

108205

From: Chan, Christina
Sent: Thursday, November 13, 2003 2:41 PM
To: Nguyen, Quang (AU1632); STIC-Biotech/ChemLib
Subject: RE: RUSH sequence search for 10/059720

Please rush. Thanks Chris

Chris Chan

TC 1600 New Hire Training Coordinator and SPE 1644
 308-3973
 CM-1, 9B19

CRFF

-----Original Message-----

From: Nguyen, Quang (AU1632)
Sent: Thursday, November 13, 2003 2:40 PM
To: Chan, Christina
Subject: RUSH sequence search for 10/059720

I would like to request for a rush sequence search for the above application, because it is an amended case due next-biweek.

Please search:

- (1) A nucleic acid encoding the amino acid sequence of SEQ ID NO:19; and
- (2) A nucleic acid of SEQ ID NO:18.

against commerical, pending and issued US Patent application databases.

I am in AU 1636, my mailbox is in CM1-11E12.

THANK YOU.

Point of Contact
 P. Sheppard
 Telephone number: (703) 308-4499

Searcher: _____
 Phone: _____
 Location: _____
 Date Picked Up: _____
 Date Completed: 11/18/03
 Searcher Prep/Review: _____
 Clerical: _____
 Online time: _____

TYPE OF SEARCH:
 NA Sequences: _____
 AA Sequences: _____
 Structures: _____
 Bibliographic: _____
 Litigation: _____
 Full text: _____
 Patent Family: _____
 Other: _____

VENDOR/COST (where applic.)
 STN: _____
 DIALOG: _____
 Questel/Orbit: _____
 DRLink: _____
 Lexis/Nexis: _____
 Sequence Sys.: _____
 WWW/Internet: _____
 Other (specify): _____

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model *

Run on: November 15, 2003, 23:00:45 ; Search time 189 Seconds
(without alignments)
4284.822 Million cell updates/sec

Title: US-10-059-720-18

Perfect score: 300

Sequence: 1 CCATGGACTACAGGACGAC.....CAAGATCCCTGATTAAAGCTT 300

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_19Jun03,*
1: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:*
2: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:*
3: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:*
4: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:*
5: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:*
6: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:*
7: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:*
8: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:*
9: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:*
10: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:*
11: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:*
12: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:*
13: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:*
14: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:*
15: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:*
16: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:*
17: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:*
18: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:*
19: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:*
20: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:*
21: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:*
22: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:*
23: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:*
24: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:*
25: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	294	98.0	294	18	AAT84348
2	265.6	88.5	281	18	AAT84347
3	173.4	57.8	318	18	AAT84346
4	157	52.3	264	18	AAT84344
5	133.4	44.5	262	18	AAT84345
6	130.8	43.6	1143	24	ABK50445
7	130.8	43.6	2084	24	ABV94327
8	130.8	43.6	2103	20	AAZ41352

9	130.8	43.6	2180	24	ABQ54467	Human ovarian anti
10	130.2	43.4	273	22	AAF79982	Nucleotide sequenc
11	124.4	41.5	1946	22	ABA44752	Human breast cell
12	124.4	41.5	1946	22	ABA55207	Human foetal liver
13	124.4	41.5	1946	22	ABA24954	Probe #3420 for ge
14	124.4	41.5	1946	22	AAK03468	Human brain expres
15	124.4	41.5	1946	22	AAK28922	Human bone marrow
16	124.4	41.5	1946	22	AAI13509	Probe #3442 for ge
17	124.4	41.5	1946	22	AAI34869	Probe #3555 used t
18	124.4	41.5	1946	22	AAI03394	Probe #3385 used t
19	124.4	41.5	1946	24	ARS03451	Human genome-deriv
20	124.4	41.5	2227	25	ABX62920	Human activated T
21	124.4	41.5	3565	19	AAV20466	Human c-fos oncoge
22	124.4	41.5	3565	21	AAF21092	Human low adenosin
23	124.4	41.5	3565	21	AAK34970	Human adenosine re
24	124.4	41.5	3565	24	ABK84488	Human cDNA differe
25	124.4	41.5	3565	24	ABK64754	Human benign prost
26	124.4	41.5	3565	24	ABL62456	Colon adenocarcino
27	124.4	41.5	3565	24	ABL68566	Kidney cancer rela
28	124.4	41.5	3565	25	ACC46762	Human COPD related
29	124.4	41.5	6210	15	AAQ63815	c-fos gene. Homo
30	124.4	41.5	6210	21	AAF21091	Human low adenosin
31	124.4	41.5	6210	21	AAA34969	Human adenosine re
32	124.4	41.5	6210	24	ABZ35089	Human gene express
33	124.4	41.5	6210	24	ABT10875	Human breast cance
34	124.4	41.5	6210	25	ACC46752	Human COPD related
35	124.4	41.5	16595	21	AAF21095	Human low adenosin
36	124.4	41.5	16595	21	AAK34973	Human adenosine re
37	124	41.3	3514	24	ABK50444	Human v-fos FBJ mu
38	123.4	41.1	642	22	ABA49894	Human breast cell
39	123.4	41.1	642	22	ABA67812	Human foetal liver
40	123.4	41.1	642	22	ABA34871	Probe #13337 for g
41	123.4	41.1	642	22	AAK16224	Human brain expres
42	123.4	41.1	642	22	AAK41968	Human bone marrow
43	123.4	41.1	642	22	AAI22734	Probe #12667 for g
44	123.4	41.1	642	22	AAI48034	Probe #16720 used
45	123.4	41.1	642	22	AAI08402	Probe #8393 used t

ALIGNMENTS

RESULT 1
AAT84348
ID AAT84348 standard; DNA; 294 BP.
XX
AC AAT84348;
XX
DT 12-NOV-1997 (first entry)
XX
DE CMV400-4HeptadFos leucine zipper protein DNA.
XX
XX DNA binding protein; RNA binding protein; amphipathic peptide;
XX acidic extension peptide; gene control; gene regulation;
XX transcription; dominant negative protein; c-Fos; cancer;
XX drug therapy; drug design; ss.
XX
XX Chimeric Homo sapiens;
XX Chimeric bacteriophage phi-10;
XX Chimeric synthetic.
XX
XX MO9705249-A2
XX
XX 13 FEB 1997.
XX
XX 31-JUL-1996; 96WO-US12590.
XX
XX 31-JUL-1996; 96US-0001654.
XX
XX 31-JUL-1996; 95US-0001654.
XX
XX 29-MAY-1996; 96US-0018496.
XX
XX (KRYL/) KRYLOV D.
XX (USSH) US DEPT HEALTH & HUMAN SERVICES.

PA (VINS/) VINSON C R.
 XX Krylov D, Vinson CR;
 XX WPI; 1997-145687/13.
 DR P-PSDB; AAW00948.
 XX
 XX New nucleic acid binding proteins - having an acidic amino acid
 PT sequence extension at the amino-terminus, to increase ability to
 PT regulate gene transcription, useful e.g. in cancer therapeutics
 XX
 XX Claim 16; Page 87; 144pp; English.
 XX
 CC This DNA sequence codes for CMV500-heptadFos leucine zipper
 CC (AAW00948), a modified Fos nucleic acid binding protein (NABP) that
 CC includes an appended 4-heptad acidic extension. Claimed NABPs such
 CC as Fos having acidic peptide extensions are capable of regulating
 CC the function of a target nucleic acid or gene to which they are
 CC bound, and act as potent dominant-negative regulators of gene
 CC transcription, cell growth and cell proliferation. They can be
 CC used in cancer therapeutics, to treat diseases caused by eukaryotic
 CC microorganisms or by viruses, and as tools for drug development,
 CC rational drug design, and drug and gene therapies. They have an
 CC extended protein interaction surface or multimerisation or
 CC dimerisation interface that increases the stability of complexes
 CC formed.
 XX
 SQ Sequence 294 BP; 96 A; 65 C; 88 G; 45 T; 0 other;
 Query Match 98.0%; Score 294; DB 18; Length 294;
 Best Local Similarity 100.0%; Pred. No. 3.1e-68;
 Matches 294; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 3 ATGGACTCAGGAGCAGATGACAGCATATGGCTAGCATGCTGGTGACGCAATG 62
 Db 1 ATGGACTCAGGAGCAGATGACAGCATATGGCTAGCATGCTGGTGACGCAATG 60
 QY 63 GTCGGGATCTGCTGCAACACGCTGCTGGAGAACTGGCCCTGAAACGAGAGTG 122
 Db 61 GTCGGGATCTGCTGCAACACGCTGCTGGAGAACTGGCCCTGAAACGAGAGTG 120
 QY 123 GAAAGAGGCGGAGAGCTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 182
 Db 121 GAAAGAGGCGGAGAGCTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 180
 QY 183 CTAGAGATGAGAGCTGCTGCTTTGAGAGCCGAGATTGCCAACCTGCTGCAAGGAGGAA 242
 Db 181 CTAGAGATGAGAGCTGCTGCTTTGAGAGCCGAGATTGCCAACCTGCTGCAAGGAGGAA 240
 QY 243 AAAGTAGATTCTGCTGCGAGCTCAGGAGCTGCTGCTGCAAGATCCCTGATTAA 296
 Db 241 AAAGTAGATTCTGCTGCGAGCTCAGGAGCTGCTGCTGCAAGATCCCTGATTAA 294
 RESULT 2
 AAT84347
 ID AAT84347 standard; DNA; 281 BP.
 XX
 AC AAT84347;
 XX
 DT 11-NOV-1997 (first entry)
 DE 4HeptadFos protein DNA.
 XX
 KW DNA binding protein; RNA binding protein; amphipathic peptide;
 KW acidic extension peptide; gene control; gene regulation;
 KW transcription; dominant negative protein; c-Fos; cancer;
 KW drug therapy; drug design; ss.
 XX
 OS Chimeric Homo sapiens;
 OS Chimeric bacteriophage phi-10;
 OS Chimeric synthetic.
 XX

PH Key Location/Qualifiers
 FT CDS 9..272
 XX /*tag= a
 XX WO9705249-A2.
 XX
 PD 13-FEB-1997.
 XX
 PF 31-JUL-1996; 96WO-US13590.
 XX
 PR 31-JUL-1996; 96US-0001654.
 PR 31-JUL-1995; 95US-0001654.
 PR 29-MAY-1996; 96US-0018496.
 XX
 PA (KRYL/) KRYLOV D.
 PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
 PA (VINS/) VINSON C R.
 XX
 PI Krylov D, Vinson CR;
 DR WPI; 1997-145687/13.
 DR P-PSDB; AAW00948.
 XX
 PT New nucleic acid binding proteins - having an acidic amino acid
 PT sequence extension at the amino-terminus, to increase ability to
 PT regulate gene transcription, useful e.g. in cancer therapeutics
 XX
 XX Claim 16; Page 86; 144pp; English.
 XX
 CC This DNA sequence codes for 4heptadFos protein (AAW00948), a
 CC modified Fos nucleic acid binding protein (NABP) that includes
 CC an appended 4-heptad acidic extension. Claimed NABPs such as
 CC Fos that have acidic peptide extensions are capable of regulating
 CC the function of a target nucleic acid or gene to which they are
 CC bound, and act as potent dominant-negative regulators of gene
 CC transcription, cell growth and cell proliferation. They can be
 CC used in cancer therapeutics, to treat diseases caused by eukaryotic
 CC microorganisms or by viruses, and as tools for drug development,
 CC rational drug design, and drug and gene therapies. They have an
 CC extended protein interaction surface or multimerisation or
 CC dimerisation interface that increases the stability of complexes
 CC formed.
 XX
 SQ Sequence 281 BP; 90 A; 63 C; 80 G; 48 T; 0 other;
 Query Match 88.5%; Score 265.6; DB 18; Length 281;
 Best Local Similarity 98.5%; Pred. No. 9.8e-61;
 Matches 268; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
 QY 25 ACAGCATATGGCTAGCATGACTGGTGACAGCAATGGTGGGATCTGACCTGGAAC 84
 Db 1 ATATACATATGGCTAGCATGACTGGTGACAGCAATGGTGGGATCTGACCTGGAAC 60
 QY 85 AACGTCTCAGGAACCTGGCCCGTGAACACGAGAGCTGGAAACAGAGCCGAGAGCTGG 144
 Db 61 AACGTCTCAGGAACCTGGCCCGTGAACACGAGAGCTGGAAACAGAGCCGAGAGCTGG 120
 QY 145 AGCAGGAAACCGCTGAACCTCGAGGCGGAGACAGACCAACTAGAGATGAGAGTCTGTT 204
 Db 121 AGCAGGAAACCGCTGAACCTCGAGGCGGAGACAGACCAACTAGAGATGAGAGTCTGTT 180
 QY 205 TCGAGCCGAGATTGCGACCTGCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 264
 Db 181 TCGAGCCGAGATTGCGACCTGCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 240
 QY 265 CTCACCGAGCTGCTGCGAGATCCCTGATTAA 296
 Db 241 CTCACCGAGCTGCTGCGAGATCCCTGATTAA 272
 RESULT 3
 AAT84346
 ID AAT84346 standard; DNA; 318 BP.

```

XX AC AAT84346;
XX DT 11-NOV-1997 (first entry)
XX DE CMV500-FosbZIP (MO) DNA.
XX KW DNA binding protein; RNA binding protein; amphipathic peptide;
XX KM acidic extension peptide; gene control; gene regulation;
XX KM transcription; dominant negative protein; c-Fos; cancer;
XX KM drug therapy; drug design; ss.
XX OS Chimeric Homo sapiens;
XX OS Chimeric bacteriophage phi 10;
XX OS Chimeric synthetic.
XX PH Key Location/Qualifiers
XX FT 1..318
XX FT CDS /*tag= a
XX FT /transl_except= (pos:1..3, aa:Met)
XX FT /transl_except= (pos:121..123, aa:Arg)
XX FN WO9705249-A2.
XX PD 13-FEB-1997.
XX PF 31-JUL-1996; 96WO-US12590.
XX PR 31-JUL-1996; 96US-0001654.
XX PR 31-JUL-1995; 95US-0001654.
XX PR 29-MAY-1996; 96US-0018496.
XX PA (KRYL/) KRYLOV D.
XX PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX PA (VINS/) VINSON C R.
XX PI Krylov D, Vinson CR;
XX DR WPI; 1997-145687/13.
XX DR P-PSDB; AAW00947.
XX PT New nucleic acid binding proteins - having an acidic amino acid
XX PT sequence extension at the amino-terminus, to increase ability to
XX PT regulate gene transcription, useful e.g. in cancer therapeutics
XX PS Claim 16; Page 85; 144pp; English.
XX CC This DNA sequence codes for CMV500-FosbZIP (WO) (AAW00947), a
XX CC modified nucleic acid binding protein (NABP). Claimed NABPs such
XX CC as Fos that have acidic peptide extensions are capable of regulating
XX CC the function of a target nucleic acid or gene to which they are
XX CC bound, and act as potent dominant-negative regulators of gene
XX CC transcription, cell growth and cell proliferation. They can be
XX CC used in cancer therapeutics, to treat diseases caused by eukaryotic
XX CC microorganisms or by viruses, and as tools for drug development.
XX CC rational drug design, and drug and gene therapies. They have an
XX CC extended protein interaction surface or multimerisation or
XX CC dimerisation interface that increases the stability of complexes
XX CC formed.
XX SQ Sequence 318 BP; 110 A; 69 C; 90 G; 49 T; 0 other;
XX
XX Query Match 57.8%; Score 173.4; DB 18; Length 318;
XX Best Local Similarity 77.4%; Pred. No. 2.4e-36;
XX Matches 247; Conservative 0; Mismatches 46; Indels 26; Gaps 2;
XX
XX 3 ATGGACTACAAGGACGACGATGACAGCATATGCGTGTAGCATGCTGGTGGACGCAATG 62
XX |||
XX 1 ATCGACTACAGGACGCGCGTGTGACAGCATATGCTGTAGCATGCTGGTGGACGCAATG 60
XX |||
XX 63 GGTCCGAGCTCTGACCTGGAAACAACTGCTGTAGGAACTGCGCCCGTGAACCAAGAGCTG 122
XX |||
XX 61 GGTCCGAGCTCTGACGAGTGGAACTGCTGTAGGAACTGCGCCCGTGAACCAAGAGCTG 119
XX |||
XX
XX QY 123 GAAAAAGAGCCCGAAGA-----GCTGGAGCAGCAAAACGCG 157
XX AC |||
XX DB 120 AAGGAAAGGATAAGATGCTGCAGCCCAATGCCGAGGAGGAGTGAATGCA 179
XX DE |||
XX DE CMV500-FosbZIP (MO) DNA.
XX KW TGAACCTCGAGCGGAGACAGACCAACTAGAGATGAGAACTCTGCTTTGCAGACCGAGAT 217
XX KM |||
XX KM TACACTCCAGCGGAGACAGACCAACTAGAGATGAGAACTCTGCTTTGCAGACCGAGAT 239
XX KM |||
XX QY 218 TGCACACTCTCTGAGAGGAGAGGAAAACTAGAGTTTCATCTGGCGAGCTCACCGACCTGC 277
XX AC |||
XX DB 240 TGCACACTCTCTGAGAGGAGAGGAAAACTAGAGTTTCATCTGGCGAGCTCACCGACCTGC 299
XX QY 278 CTGCAAGATCCCTGATTAA 296
XX DB 300 CTGCAAGATCCCTGATTAA 318
XX
XX RESULT 4
XX AAT84344
XX ID AAT84344 standard; DNA; 264 BP.
XX AC AAT84344;
XX DT 11-NOV-1997 (first entry)
XX DE DNA encoding CMV500-4heptadCREB (New4hepCREB).
XX KW DNA binding protein; RNA binding protein; amphipathic peptide;
XX KM acidic extension peptide; gene control; gene regulation;
XX KM transcription; dominant negative protein; CREB; cancer;
XX KM drug therapy; drug design; CMV; ss.
XX OS Chimeric Homo sapiens;
XX OS Chimeric cytomegalovirus;
XX OS Chimeric synthetic.
XX PH Key Location/Qualifiers
XX FT 1..255
XX FT CDS /*tag= a
XX FT /transl_except= (pos:229..231, aa:Glu)
XX
XX WO9705249-A2.
XX PD 13-FEB-1997.
XX PF 31-JUL-1996; 96WO-US12590.
XX PR 31-JUL-1996; 96US-0001654.
XX PR 31-JUL-1995; 95US-0001654.
XX PR 29-MAY-1996; 96US-0018496.
XX PA (KRYL/) KRYLOV D.
XX PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX PA (VINS/) VINSON C R.
XX PI Krylov D, Vinson CR;
XX DR WPI; 1997-145687/13.
XX DR P-PSDB; AAW00945.
XX PT New nucleic acid binding proteins - having an acidic amino acid
XX PT sequence extension at the amino-terminus, to increase ability to
XX PT regulate gene transcription, useful e.g. in cancer therapeutics
XX PS Claim 16; Page 83; 144pp; English.
XX CC This DNA sequence codes for CMV500-4heptadCREB (AAW00945), or
XX CC New4hepCREB, which has a 4heptad appended acidic extension.
XX CC Claimed nucleic acid binding proteins (NABPs) such as CREB that
XX CC have acidic peptide extensions are capable of regulating the
XX CC function of a target nucleic acid or gene to which they are bound,
XX CC and act as potent dominant-negative regulators of gene

```

CC transcription, cell growth and cell proliferation. They can be
 CC used in cancer therapeutics, to treat diseases caused by eukaryotic
 CC microorganisms or by viruses, and as tools for drug development,
 CC rational drug design, and drug and gene therapies. They have an
 CC extended protein interaction surface or multimerisation or
 CC dimerisation interface that increases the stability of complexes
 CC formed.
 XX
 SQ Sequence 264 BP; 93 A; 52 C; 76 G; 43 T; 0 other;
 Query Match 52.3%; Score 157; DB 18; Length 264;
 Best Local Similarity 97.0%; Pred. No. 4.9e-32;
 Matches 160; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
 QY 3 ATGGACTACAAGGACGACGATGACAAAGCATATGGCTAGCATGACTGTGGACACCAATG 62
 DB 1 ATGGACTACAAGGACGACGATGACAAAGCATATGGCTAGCATGACTGTGGACACCAATG 60
 QY 63 GGTGGGATCTGACCTGGAAACAGCTGCTGAGGAACCTGGCCCGTGAACCAAGAGAGCTG 122
 DB 61 GGTGGGATCTGACCTGGAAACAGCTGCTGAGGAACCTGGCCCGTGAACCAAGAGAGCTG 120
 QY 123 GAAAAAGGCGGAGAGCTGGAGCAGGAAACGCTGAACCTGAG 167
 DB 121 GAAAAAGGCGGAGAGCTGGAGCAGGAAACGCTGAG 165
 RESULT 5
 AAT84345
 ID AAT84345 standard; DNA; 262 BP.
 AC AAT84345;
 XX
 DT 11-NOV-1997 (first entry)
 XX
 DE Human c-Fos protein DNA.
 XX
 KW DNA binding protein; RNA binding protein; amphipathic peptide;
 KW acidic extension peptide; gene control; gene regulation;
 KW transcription; dominant negative protein; c-Fos; cancer;
 KW drug therapy; drug design; ss.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT CDS 2..253
 FT /*tag= a
 FT /transl_except= (pos:1..3, aa:Pro)
 FT /transl_except= (pos:4..6, aa:Asp)
 XX
 PN W09705249-A2.
 XX
 PD 13-FEB-1997.
 XX
 PF 31-JUL-1996; 96WO-US12590.
 XX
 PR 31-JUL-1996; 96US-0001654.
 PR 31-JUL-1995; 95US-0001654.
 PR 29-MAY-1996; 96US-0018496.
 XX
 XX (KRYL/) KRYLOV D.
 PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
 PA (VINS/) VINSON C R.
 XX
 PI Krylov D, Vinson CR;
 XX
 DR WPI; 1997-145687/13.
 DR P-PSDB; AAW00946.
 XX
 XX New nucleic acid binding proteins - having an acidic amino acid
 FT sequence extension at the amino-terminus, to increase ability to
 FT regulate gene transcription, useful e.g. in cancer therapeutics
 XX.

PS Claim 16; Page 84; 144pp; English.
 CC
 CC This DNA sequence codes for the human c-Fos (AAW00946), a nucleic
 CC acid binding protein (NABP). Claimed NABPs such as c-Fos that
 CC have acidic peptide extensions are capable of regulating the
 CC function of a target nucleic acid or gene to which they are bound,
 CC and act as potent dominant-negative regulators of gene
 CC transcription, cell growth and cell proliferation. They can be
 CC used in cancer therapeutics, to treat diseases caused by eukaryotic
 CC microorganisms or by viruses, and as tools for drug development,
 CC rational drug design, and drug and gene therapies. They have an
 CC extended protein interaction surface or multimerisation or
 CC dimerisation interface that increases the stability of complexes
 CC formed.
 XX
 SQ Sequence 262 BP; 90 A; 58 C; 72 G; 42 T; 0 other;
 Query Match 44.5%; Score 133.4; DB 18; Length 262;
 Best Local Similarity 92.7%; Pred. No. 8.5e-26;
 Matches 140; Conservative 0; Mismatches 11; Indels 0; Gaps 0;
 QY 146 GCGGAAACGCTGAACCTCGAGCGGAGACAGACCAACTAGAAGTCTGCTTT 205
 DB 103 GGAGCTGACTGATACACTCCAGCGGAGACAGACCAACTAGAAGTCTGCTTT 162
 QY 206 GCGAGCCGAGATTGCCAACCCTGCTGAGGAGGAGGAAACTAGATTCTATCTGGCAGC 265
 DB 163 GCGAGCCGAGATTGCCAACCCTGCTGAGGAGGAGGAAACTAGATTCTATCTGGCAGC 222
 QY 266 TCACCGACCTGCTGCAAGATCCCTGATTA 296
 DB 223 TCACCGACCTGCTGCAAGATCCCTGATTA 253
 RESULT 6
 ABK50445
 ID ABK50445 standard; cDNA; 1143 BP.
 AC ABK50445;
 XX
 DT 13-AUG-2002 (first entry)
 XX
 DE Human v-fos FBJ murine osteosarcoma viral oncogene homologue (FOS) cDNA.
 XX
 KW Human; v-fos FBJ murine osteosarcoma viral oncogene homologue; FOS; gene;
 KW cytostatic; gene therapy; single nucleotide polymorphism; haplotyping;
 KW haplotype pair; developmental bone disorder; cancer; tumour; ss;
 KW chromosome 14q21-q31.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT CDS 1..1143
 FT /*tag= a
 FT /product= "Human FOS protein"
 FT replace(189,A)
 FT /*tag= b
 FT /standard_name= "Single nucleotide polymorphism"
 FT replace(252,C)
 FT /*tag= c
 FT /standard_name= "Single nucleotide polymorphism"
 FT replace(528,C)
 FT /*tag= d
 FT /standard_name= "Single nucleotide polymorphism"
 FT replace(699,A)
 FT /*tag= e
 FT /standard_name= "Single nucleotide polymorphism"
 FT replace(846,G)
 FT /*tag= f
 FT /standard_name= "Single nucleotide polymorphism"
 XX
 PN W0200232931-A2.
 XX

XX 19-JAN-2000 (first entry)
 XX Human normal uterus tissue derived cDNA 28.
 DE Human; uterus; cancer; treatment; anticancer; cytostatic; gene therapy;
 KW EST; expressed sequence tag; ss.
 KW Homo sapiens.
 OS DE19817946-A1.
 XX 21-OCT-1999.
 XX 17-APR-1998; 98DE-1017946.
 XX 17-APR-1998; 98DE-1017946.
 PA (META-) METAGEN GES GENOMFORSCHUNG MBH.
 XX Rosenthal A, Specht T, Hinzmann B, Schmitt A, Pilarsky C, Dahl E;
 DR WPI; 1999-591956/51.
 XX New nucleic acid sequences expressed in normal uterine tissues, and
 PT derived polypeptides, for treatment of uterine cancer and
 PT identification of therapeutic agents -
 XX Claim 3; Page 94; 154pp; German.
 XX This invention describes novel cDNA sequences (A) highly expressed in
 CC normal uterine tissue which can have anticancer and cytostatic activity
 CC and can be used for gene therapy. (A) are used (i) for recombinant
 CC expression of polypeptides (B) and (ii) to isolate complete genes.
 CC (B) are used (i) to identify agents suitable for treatment of uterine
 CC cancer; (ii) directly for treating this form of cancer (including
 CC expression from gene therapy vectors) and (iii) for generation of
 CC specific antibodies. (A) are identified by assembling ESTs (expressed
 CC sequence tags) from a particular tissue type before comparison of
 CC expression patterns. This allows a significantly longer fragment of the
 CC gene to be revealed, so should reduce the number of failures associated
 CC with the fact that ESTs from different libraries may represent
 CC different parts of the same unknown gene, distorting the estimated
 CC frequency of occurrence in a particular tissue. AM241325-241385
 CC represent the human uterine tissue derived cDNA fragments of the
 CC invention which encode the protein fragments represented in
 CC AAY59838-Y59892.
 XX Sequence 2103 BP; 492 A; 573 C; 525 G; 513 T; 0 other;
 QY Query Match 43.6%; Score 130.8; DB 20; Length 2103;
 Db Best Local Similarity 92.0%; Pred. No. 7.4e-25;
 Mismatches 138; Conservative 0; Mismatches 12; Indels 0; Gaps 0;
 QY 146 GCAGGAAACGCTGACTCGAGCGGAGACAGACCACTAGAGATGAGAGTCTGCTTT 205
 Db 588 GGGCTGACTGATACACTCGAGCGGAGACAGACCACTAGAGATGAGAGTCTGCTTT 647
 QY 206 GCAGACCGAGATTGCCAATCTGCTGAGGAGAGGAGAACTAGATTCCTCGGAGC 265
 Db 648 GCAGACCGAGATTGCCAATCTGCTGAGGAGAGGAGAACTAGATTCCTCGGAGC 707
 QY 266 TCACCGACTGCTGCAAGTCCCTGATTA 295
 Db 708 TCACCGACTGCTGCAAGTCCCTGATTA 737

RESULT 9
 ABQ54467
 ID ABQ54467 standard; cDNA; 2180 BP.
 XX AC
 XX ABQ54467;
 XX

22-AUG-2002 (first entry)
 Human ovarian antigen HTHDV01 cDNA, SEQ ID NO:347.
 KW Human; ovarian antigen; ovary; ovarian; breast; cancer; tumour;
 KW ovarian cancer; breast cancer; tumour; reproductive system disorder;
 KW infertility; pregnancy disorder; anovulation; polycystic ovary syndrome;
 KW PCOS; ovarian cyst; dysmenorrhoea; endocrine disorder; infection;
 KW inflammatory condition; immune disorder; blood disorder;
 KW cardiovascular disorder; respiratory disorder; neurological disorder;
 KW gastrointestinal disorder; urinary system disorder; drug screening;
 KW gene therapy; chromosome mapping; forensic analysis;
 KW antibody preparation; cytostatic; immunomodulatory; neuroprotective;
 KW antiinflammatory; gynaecological; reproductive; chromosome 14q24.3;
 KW gene; ss.
 XX Homo sapiens.
 XX WO200200677-A1.
 XX 03-JAN-2002.
 XX 07-JUN-2001; 2001WO-US18569.
 XX 07-JUN-2000; 2000US-209467P.
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX Birse CB, Rosen CA;
 PI WPI; 2002-147878/19.
 DR P-PSDB; ABP41390.
 XX Isolated nucleic acid molecules encoding novel ovarian polypeptides,
 PT useful in the prevention, treatment and diagnosis of cancer (e.g.
 PT ovarian cancer), immune disorders, cardiovascular disorders and
 PT neurological diseases -
 XX Claim 1; SEQ ID NO 347; 2922pp; English.
 XX The invention relates to 2175 novel human ovarian antigens (ABP41054-
 CC ABP43228) and to cDNAs encoding them (ABQ54131-ABQ56305), and also
 CC encompasses polypeptides 90% identical and polynucleotides 95% identical
 CC to the sequences of the invention. The invention additionally relates to
 CC recombinant vectors and host cells comprising human ovarian antigen
 CC polynucleotides, antibodies against human ovarian antigens, and the use
 CC of ovarian antigen polynucleotides and polypeptides in diagnosing,
 CC treating, prognosing or preventing various ovary and/or breast-related
 CC disorders. Such conditions include ovarian cancer and breast cancer, and
 CC metastatic tumours of ovarian or breast origin, reproductive system
 CC disorders (e.g., infertility, disorders of pregnancy, anovulation,
 CC polycystic ovary syndrome, ovarian cysts, and dysmenorrhoea), endocrine
 CC disorders, infections (e.g., chlamydia, HIV, toxoplasmosis, and toxic
 CC shock syndrome), inflammatory conditions (e.g., mastitis, oophoritis and
 CC vaginitis), immune disorders (e.g., congenital and acquired
 CC immunodeficiencies, autoimmune oophoritis, systemic lupus erythematosus),
 CC blood-related disorders (e.g., anaemia), cardiovascular disorders,
 CC respiratory disorders, neurological disorders, gastrointestinal disorders
 CC and urinary system disorders. Ovarian antigen polypeptides and
 CC polynucleotides may also be used in screening for compounds which
 CC modulate ovarian antigen expression or activity. The polynucleotides may
 CC further be used for gene therapy, chromosome mapping, in the
 CC identification of individuals and in forensic analysis, and the
 CC polypeptides may be used as food additives or to prepare antibodies
 CC useful in disease diagnosis, drug targeting and phenotyping. The present
 CC sequence represents cDNA encoding a human ovarian antigen of the
 CC invention.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX Sequence 2180 BP; 513 A; 596 C; 549 G; 519 T; 3 other;

Query Match 43.6%; Score 130.8; DB 24; Length 2180;
Best Local Similarity 92.0%; Pred. No. 7.4e-25;
Matches 138; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

QY 146 GCAGGAAACGCTGAACCTCGAGCGGAGACAGACCAACTAGAGATGAGAACTCTGCTTT 205
DB 629 GGAGCTGACTGATACACTCCAGCGGAGACAGACCAACTAGAGATGAGAACTCTGCTTT 688
QY 206 GCAGACCGGAGATGCCAAGCTCTGAGGAGAGAGAACTAGAGTTCACTCTGGCAGC 265
DB 689 GCAGACCGGAGATGCCAAGCTCTGAGGAGAGAGAACTAGAGTTCACTCTGGCAGC 748
QY 266 TCACCGACCTGCTGCAAGATCCCTGATTA 295
DB 749 TCACCGACCTGCTGCAAGATCCCTGATTA 778

RESULT 10

AAF79982/c

ID AAF79982 standard; DNA; 273 BP.

XX

AC AAF79982;

XX

DT 11-JUN-2001 (first entry)

XX

DE Nucleotide sequence of a human genetic marker for toxicity.

XX

KW Genetic marker; toxicity; cellular signalling pathway; polymorphism; ss.

XX

OS Homo sapiens.

XX

FN WO200120029-A2.

XX

PD 22-MAR-2001.

XX

PF 12-SEP-2000; 2000WO-FR02503.

XX

PR 13-SEP-1999; 99FR-0011405.

XX

PA (EXON-) EXONHIT THERAPEUTICS SA.

XX

PI Tocque B, Bracco L, Schweighoffer F;

XX

DR WPI; 2001-244821/25.

XX

PT Analysing the toxic potential of test compounds, for e.g. screening

XX

PS Claim 35; Page 61; 68pp; French.

XX

CC AAF79967-AAF80003 represents genetic markers of toxicity. The
CC specification describes a method for analysing the toxic potential
CC of a test compound. The method comprises hybridising nucleic acids
CC from cells treated with the test compound and the present markers.
CC These markers correspond to genetic events characteristic of
CC deregulation of cellular signalling pathways. The method is used to
CC identify the toxic potential of compounds (particularly human or
CC veterinary pharmaceuticals or plant protection agents) and to evaluate
CC the response and/or sensitivity of subjects to a particular compound,
CC from the presence of polymorphisms or other mutations in particular
CC genes.

XX

SQ Sequence 273 BP; 52 A; 74 C; 79 G; 68 T; 0 other;

XX

Query Match 43.4%; Score 130.2; DB 22; Length 273;
Best Local Similarity 97.8%; Pred. No. 6e-25;
Matches 132; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 161 ACTCGAGCGGAGACAGACCAACTAGAGATGAGAACTCTGCTTTCGACCGAGATTGC 220
DB 237 ACTCGAGCGGAGACAGACCAACTAGAGATGAGAACTCTGCTTTCGACCGAGATTGC 178

QY 221 CAACCTGCTGAAGGAGAGGAAAAAAGTAGAGTTCTCTGGCAGCTCACCGACCTGCCTG 280
DB 177 CAACCTGCTGAAGGAGAGGAAAAAAGTAGAGTTCTCTGGCAGCTCACCGACCTGCCTG 118
QY 281 CAAGATCCCTGATTA 295
DB 117 CAAGATCCCTGATTA 103

RESULT 11

ABA44752

ID ABA44752 standard; DNA; 1946 BP.

XX

AC ABA44752;

XX

DT 01-FEB-2002 (first entry)

XX

DE Human breast cell single exon nucleic acid probe #3447.

XX

KW Human; microarray; single exon probe; gene expression; breast;

XX

KW disease; cancer; ss.

XX

OS Homo sapiens.

XX

FN WO200157271-A2.

XX

PD 09-AUG-2001.

XX

PF 30-JAN-2001; 2001WO-US00662.

XX

PR 04-FEB-2000; 2000US-0180312.

XX

PR 26-MAY-2000; 2000US-0207456.

XX

PR 30-JUN-2000; 2000US-0608408.

XX

PR 03-AUG-2000; 2000US-0632366.

XX

PR 21-SEP-2000; 2000US-0234687.

XX

PR 27-SEP-2000; 2000US-0236359.

XX

PR 04-OCT-2000; 2000GB-0024263.

XX

PA (MOLE-) MOLECULAR DYNAMICS INC.

XX

PI Penn SG, Hanzel DK, Chen W, Rank DR;

XX

DR WPI; 2001-496933/54.

XX

PT New spatially-addressable set of single exon nucleic acid probes,
PT useful for measuring gene expression in sample derived from human
PT breast, comprises number of single exon nucleic acid probes -
XX

PS Claim 1; SEQ ID NO 3447; 327pp + sequence listing; English.

XX

CC The invention relates to a spatially-addressable set of single exon
CC nucleic acid probes for measuring gene expression in a sample derived
CC from human breast and BT 474 cells. The method involves contacting
CC the probes with a collection of detectably labelled nucleic acids
CC derived from mRNA of human breast and then measuring the label
CC bound to each probe of the microarray. The probes are useful for
CC verifying the expression of regions of genomic DNA predicted to
CC encode proteins. They are useful for gene discovery, and for
CC determining predisposition and/or prognosing breast disease. Gene
CC expression analysis is useful for assessing the toxicity of chemical
CC agents on cells. The microarray of this invention presents a far greater
CC diversity of probes for measuring gene expression, with far less bias
CC than expressed sequence tag microarrays. The method is suitable for
CC rapid production of functional information from genomic sequence. The
CC present sequence is a single exon nucleic acid probe of the invention.
CC Note: The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at fip.wipo.int/pub/published_pct_sequences.

XX

SQ Sequence 1946 BP; 482 A; 440 C; 476 G; 548 T; 0 other;

XX

Query Match 41.5%; Score 124.4; DB 22; Length 1946;

Best Local Similarity 99.2%; Pred. No. 3.5e-23;

Matches 125; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 170 GGAGACAGACCACTAGAGATGAGAGTCTGTTTGCAGACCGAGATTGCCAAGCTGCT 229
DB 313 GGAGACAGACCACTAGAGATGAGAGTCTGTTTGCAGACCGAGATTGCCAAGCTGCT 372
QY 230 GAAGGAGAGGAAAACTAGATTTCATCTGCGAGCTCACCGACCTGCTGCAAGATCCC 289
DB 373 GAAGGAGAGGAAAACTAGATTTCATCTGCGAGCTCACCGACCTGCTGCAAGATCCC 432
QY 290 TGATTA 295
DB 433 TGATGA 438

RESULT 12

ABA55207
ID ABA55207 standard; DNA; 1946 BP.

AC ABA55207;

DT 01-FEB-2002 (first entry)
DE Human foetal liver single exon nucleic acid probe #3512.

KW Human; foetal liver; gene expression; single exon nucleic acid probe; ss.

OS Homo sapiens.
PN WO200157277-A2.
PD 09-AUG-2001.

PF 30-JAN-2001; 2001WO-US00669.

PR 04-FEB-2000; 2000US-0180312.

PR 26-MAY-2000; 2000US-0207456.

PR 30-JUN-2000; 2000US-0608408.

PR 03-AUG-2000; 2000US-0632366.

PR 21-SEP-2000; 2000US-0234687.

PR 27-SEP-2000; 2000US-0236359.

PR 04-OCT-2000; 2000GB-0024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

Penn SG, Hanzel DK, Chen W, Rank DR;

WPI; 2001-483447/52.

Human genome-derived single exon nucleic acid probes useful for
analyzing gene expression in human fetal liver -
Claim 1; SEQ ID NO 3512; 639pp + sequence listing; English.

The invention relates to a single exon nucleic acid probe for
measuring human gene expression in a sample derived from human foetal
liver. The single exon nucleic acid probes may be used for predicting,
measuring and displaying gene expression in samples derived from human
fetal liver. The present sequence is a single exon nucleic acid
probe of the invention.
CC Note: The sequence data for this patent did not form part of the
printed specification, but was obtained in electronic format directly
from WIPO at ftp.wipo.int/pub/published_pct_sequences.

Sequence 1946 BP; 482 A; 440 C; 476 G; 548 T; 0 other;
Query Match 41.5%; Score 124.4; DB 22; Length 1946;
Best Local Similarity 99.2%; Pred. No. 3.5e-23;
Matches 125; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 170 GGAGACAGACCACTAGAGATGAGAGTCTGTTTGCAGACCGAGATTGCCAAGCTGCT 229
DB 313 GGAGACAGACCACTAGAGATGAGAGTCTGTTTGCAGACCGAGATTGCCAAGCTGCT 372

QY 230 GAAGGAGAGGAAAACTAGATTTCATCTGCGAGCTCACCGACCTGCTGCAAGATCCC 289
DB 373 GAAGGAGAGGAAAACTAGATTTCATCTGCGAGCTCACCGACCTGCTGCAAGATCCC 432
QY 290 TGATTA 295
DB 433 TGATGA 438

RESULT 13

ABA24954

ID ABA24954 standard; DNA; 1946 BP.

AC ABA24954;

DT 23-JAN-2002 (first entry)
DE Probe #3420 for gene expression analysis in human heart cell sample.

KW Human; gene expression; heart; microarray; vascular system; probe;
cardiovascular disease; hypertension; cardiac arrhythmia;
congenital heart disease; ss.

OS Homo sapiens.

PN WO200157274-A2.

PD 09-AUG-2001.

PF 30-JAN-2001; 2001WO-US00666.

PR 04-FEB-2000; 2000US-0180312.

PR 26-MAY-2000; 2000US-0207456.

PR 30-JUN-2000; 2000US-0608408.

PR 03-AUG-2000; 2000US-0632366.

PR 21-SEP-2000; 2000US-0234687.

PR 27-SEP-2000; 2000US-0236359.

PR 04-OCT-2000; 2000GB-0024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

Penn SG, Hanzel DK, Chen W, Rank DR;

WPI; 2001-488899/53.

Single exon nucleic acid probes for analyzing gene expression in human
hearts -
Claim 1; SEQ ID No 3420; 530pp; English.
The present invention relates to single exon nucleic acid probes for
measuring human gene expression in a sample derived from human heart. The
present sequence is one such probe. The probes may be used for
predicting, measuring and displaying gene expression in samples derived
from the human heart via microarrays. By measuring gene expression, the
probes are useful for predicting, diagnosing, grading, staging,
monitoring and prognosing diseases of the human heart and vascular system
e.g. cardiovascular disease, hypertension, cardiac arrhythmias and
congenital heart disease.
CC Note: The sequence data for this patent did not form part of the printed
specification, but was obtained in electronic format directly from WIPO
at ftp.wipo.int/pub/published_pct_sequences.

Sequence 1946 BP; 482 A; 440 C; 476 G; 548 T; 0 other;
Query Match 41.5%; Score 124.4; DB 22; Length 1946;
Best Local Similarity 99.2%; Pred. No. 3.5e-23;
Matches 125; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 170 GGAGACAGACCACTAGAGATGAGAGTCTGTTTGCAGACCGAGATTGCCAAGCTGCT 229
DB 313 GGAGACAGACCACTAGAGATGAGAGTCTGTTTGCAGACCGAGATTGCCAAGCTGCT 372

QY 230 GAAGGAGAGGAAAACTAGAGTTTCATCTGGCAGCTCACCAGCTGCTGCTGCAGATCCC 289
 Db |||||
 373 GAAGGAGAGGAAAACTAGAGTTTCATCTGGCAGCTCACCAGCTGCTGCTGCAGATCCC 432

QY 290 TGATTA 295
 Db |||||
 433 TGATGA 438

RESULT 14

AAK03468
 ID AAK03468 standard; DNA; 1946 BP.

XX AC
 AC AAK03468;

XX DT 05-NOV-2001 (first entry)

XX DE Human brain expressed single exon probe SEQ ID NO: 3459.

XX KW Human; brain expressed exon; gene expression analysis; probe;
 KW microarray; Alzheimer's disease; multiple sclerosis; schizophrenia;
 KW epilepsy; cancer; ss.

XX OS Homo sapiens.

XX PN WO200157275-A2.

XX PD 09-AUG-2001.

XX PF 30-JAN-2001; 2001WO-US00667.

XX PR 04-FEB-2000; 2000US-0180312.

XX PR 26-MAY-2000; 2000US-0207456.

XX PR 30-JUN-2000; 2000US-0608408.

XX PR 03-AUG-2000; 2000US-0632366.

XX PR 21-SEP-2000; 2000US-0234687.

XX PR 27-SEP-2000; 2000US-0236359.

XX PR 04-OCT-2000; 2000GB-0024263.

XX PA (MOLE-) MOLECULAR DYNAMICS INC.

XX PI Penn SG, Hanzel DK, Chen W, Rank DR;

XX PS WPI; 2001-483446/S2.

XX PT Single exon nucleic acid probes for analyzing gene expression in human
 PT brains -

XX PS Example 4; SEQ ID NO: 3459; 650pp + Sequence Listing; English.

XX CC The present invention provides a number of single exon nucleic acid
 CC probes which are derived from genomic sequences expressed in the human
 CC brain. They can be used to measure gene expression in brain cell samples,
 CC which may enable the diagnosis and improved treatment of nervous system
 CC diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,
 CC epilepsy and cancers. The present sequence is one of the probes of the
 CC invention.

XX SQ Sequence 1946 BP; 482 A; 440 C; 476 G; 548 T; 0 other;

Query Match 41.5%; Score 124.4; DB 22; Length 1946;

Best Local Similarity 99.2%; Pred. No. 3.5e-23;

Matches 125; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 170 GGAGACAGACCAACTAGAGATGAGAGTCTCTTTGAGACCGAGATTGCCAAGTCTCT 229

Db |||||

313 GGAGACAGACCAACTAGAGATGAGAGTCTCTTTGAGACCGAGATTGCCAAGTCTCT 372

QY 230 GAAGGAGAGGAAAACTAGAGTTTCATCTGGCAGCTCACCAGCTGCTGCTGCAGATCCC 289

Db |||||

373 GAAGGAGAGGAAAACTAGAGTTTCATCTGGCAGCTCACCAGCTGCTGCTGCAGATCCC 432

QY 290 TGATTA 295
 Db |||||
 433 TGATGA 438

RESULT 15

AAK28922
 ID AAK28922 standard; DNA; 1946 BP.

XX AC
 AC AAK28922;

XX DT 06-NOV-2001 (first entry)

XX DE Human bone marrow expressed single exon probe SEQ ID NO: 3479.

XX KW Human; bone marrow expressed exon; gene expression analysis; probe;
 KW microarray; cancer; leukaemia; lymphoma; myeloma; ss.

XX OS Homo sapiens.

XX PN WO200157276-A2.

XX PD 09-AUG-2001.

XX PF 30-JAN-2001; 2001WO-US00668.

XX PR 04-FEB-2000; 2000US-0180312.

XX PR 26-MAY-2000; 2000US-0207456.

XX PR 30-JUN-2000; 2000US-0608408.

XX PR 03-AUG-2000; 2000US-0632366.

XX PR 21-SEP-2000; 2000US-0234687.

XX PR 27-SEP-2000; 2000US-0236359.

XX PR 04-OCT-2000; 2000GB-0024263.

XX PA (MOLE-) MOLECULAR DYNAMICS INC.

XX PI Penn SG, Hanzel DK, Chen W, Rank DR;

XX PS WPI; 2001-488900/S3.

XX PT Human genome-derived single exon nucleic acid probes useful for
 PT analyzing gene expression in human bone marrow -

XX PS Example 4; SEQ ID NO: 3479; 658pp + Sequence Listing; English.

XX CC The present invention provides a number of single exon nucleic acid
 CC probes which are derived from genomic sequences expressed in the human
 CC bone marrow. They can be used to measure gene expression in bone marrow
 CC samples, which may enable the improved diagnosis and treatment of cancers
 CC such as lymphoma, leukaemia and myeloma. The present sequence is one of
 CC the probes of the invention.

XX SQ Sequence 1946 BP; 482 A; 440 C; 476 G; 548 T; 0 other;

Query Match 41.5%; Score 124.4; DB 22; Length 1946;

Best Local Similarity 99.2%; Pred. No. 3.5e-23;

Matches 125; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 170 GGAGACAGACCAACTAGAGATGAGAGTCTCTTTGAGACCGAGATTGCCAAGTCTCT 229

Db |||||

313 GGAGACAGACCAACTAGAGATGAGAGTCTCTTTGAGACCGAGATTGCCAAGTCTCT 372

QY 230 GAAGGAGAGGAAAACTAGAGTTTCATCTGGCAGCTCACCAGCTGCTGCTGCAGATCCC 289

Db |||||

373 GAAGGAGAGGAAAACTAGAGTTTCATCTGGCAGCTCACCAGCTGCTGCTGCAGATCCC 432

QY 290 TGATTA 295

Db |||||

433 TGATGA 438

Search completed: November 16, 2003, 00:10:34

Job time : 190 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 16, 2003, 00:02:30 ; Search time 1315 Seconds
(without alignments)
5544.747 Million cell updates/sec

Title: US-10-059-720-18

Perfect score: 300

Sequence: 1 CCATGACTACAGGACGAC.....CAAGATCCCTGATTAGCTT 300

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST.*

1: em_estba.*

2: em_esthum.*

3: em_estin.*

4: em_estmu.*

5: em_estov.*

6: em_estpl.*

7: em_estro.*

8: em_esti.*

9: gb_esti.*

10: gb_est2.*

11: gb_est3.*

12: gb_est4.*

13: gb_est5.*

14: gb_est6.*

15: em_estfun.*

16: em_estom.*

17: em_gss_hum.*

18: em_gss_inv.*

19: em_gss_pln.*

20: em_gss_vrt.*

21: em_gss_fun.*

22: em_gss_man.*

23: em_gss_mus.*

24: em_gss_pro.*

25: em_gss_rod.*

26: em_gss_phg.*

27: em_gss_vri.*

28: gb_gss1.*

29: gb_gss2.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	132.2	44.1	439	10	BF724668
2	132.2	44.1	590	14	CB266656
3	132.2	44.1	597	14	CB269390
4	132.2	44.1	623	14	CA407561

5	130.8	43.6	288	14	CA388732
6	130.8	43.6	326	13	BQ082970
7	130.8	43.6	337	10	AW881177
8	130.8	43.6	386	10	BF298872
9	130.8	43.6	396	12	BI019270
10	130.8	43.6	519	12	BM111302
11	130.8	43.6	550	9	AL703189
12	130.8	43.6	551	13	BQ894927
13	130.8	43.6	553	10	BE840522
14	130.8	43.6	557	10	BE840523
15	130.8	43.6	565	14	CA406906
16	130.8	43.6	581	12	BM795093
17	130.8	43.6	614	13	BQ694699
18	130.8	43.6	615	12	BQ324287
19	130.8	43.6	635	13	BQ894763
20	130.8	43.6	637	12	BM818004
21	130.8	43.6	641	12	BM352352
22	130.8	43.6	650	13	BQ693998
23	130.8	43.6	654	12	BM820048
24	130.8	43.6	667	12	BI481329
25	130.8	43.6	673	12	BM312542
26	130.8	43.6	691	12	BM868586
27	130.8	43.6	713	12	BQ925313
28	130.8	43.6	755	14	CA777910
29	130.8	43.6	778	14	CA772039
30	130.8	43.6	789	10	BG695872
31	130.8	43.6	794	12	BI908747
32	130.8	43.6	849	14	CD107570
33	130.8	43.6	895	13	BQ171306
34	130.8	43.6	916	13	EX415775
35	130.8	43.6	923	10	BG742465
36	130.8	43.6	961	13	BQ892958
37	130.8	43.6	975	13	BQ882100
38	130.8	43.6	976	12	BM545725
39	130.8	43.6	1129	13	BQ723263
40	130.8	43.6	1201	13	EX360805
41	130.4	43.5	302	12	BG950080
42	130.4	43.5	428	9	AU297079
43	130.4	43.5	467	13	BQ130121
44	130.4	43.5	831	10	BF337994
45	130	43.3	460	14	N36936

ALIGNMENTS

RESULT 1

BF724668

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

BF724668 439 bp mRNA linear EST 05-JAN-2001
bx07e01.y1 Human Iris cDNA (Un-normalized, unamplified): BX Homo
sapiens cDNA clone bx07e01.5', mRNA sequence.

BF724668

BF724668.1 GI:12040577

EST.

Homo sapiens (human)

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

Wistow,G.J., Bernstein,S., Behal,A. and Smith,D.

NEBANK: EST analysis and bioinformatics for ocular genomics

Invest. Ophthalmol. Vis. Sci. 41, (2000) In press

Contact: Wistow G

Section on Molecular Structure and Function

National Eye Institute

6/331, NIH, Bethesda, MD 20892-2740, USA

Tel: 301 402 3452

Fax: 301 496 0078

Email: graeme@helix.nih.gov

Plate: 07 row: e column: 01

Seq primer: M13RPI reverse primer (ABI).

Location/Qualifiers

1. .439

source

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
REFERENCE 1 (bases 1 to 386)
AUTHORS Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R., Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F., Goldman, G.H., Carvalho, A.F., Matsukuma, A., Baia, G.S., Simpson, D.H., Brunstein, A., de Oliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and Simpson, A.J.
TITLE Shotgun sequencing of the human transcriptome with ORF expressed sequence tags
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE 20202663
PUBMED 10737800
COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?cl=MR3&t2=MR3-RN0052-261200-002-d11&t3=2000-12-26&t4=1)
Seq primer: puc 18 forward
High quality sequence start: 14
High quality sequence stop: 386.
FEATURES
source
1..386
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="Adult"
/clone_lib="RN0052"
/note="Organ: head normal; Vector: puc18; Site: 1: Smal; Site 2: Site 2: Smal; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the puc 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."
BASE COUNT 77 a 100 c 114 g 93 t 2 others
ORIGIN
Query Match 43.6%; Score 130.8; DB 10; Length 386;
Best Local Similarity 92.0%; Pred. No. 1.3e-21;
Matches 138; Conservative 0; Mismatches 12; Indels 0; Gaps 0;
QY 146 GCAGGAACGCTGAACCTGAGCGGAGACAGACCACTAGAGATGAGAGTCTGCTTT 205
DB 375 GGAGCTGACTGATACACTCCAGCGGAGACAGACCACTAGAGATGAGAGTCTGCTTT 316
QY 206 GCAGACCGAGATTGCCAATCTGCTGAAGGAGAGAGAACTAGAGTTCATCTTGGCAGC 265
DB 315 GCAGACCGAGATTGCCAATCTGCTGAAGGAGAGAGAACTAGAGTTCATCTTGGCAGC 256
QY 266 TCACCGACCTGCTGCAAGATCCCTGATTA 295
DB 255 TCACCGACCTGCTGCAAGATCCCTGATGA 226
RESULT 9
BI019270
LOCUS BI019270 396 bp mRNA linear EST 14-JUN-2001
DEFINITION PM3-MT0200-060101-004-g12 MT0200 Homo sapiens cDNA, mRNA sequence.
ACCESSION BI019270
VERSION BI019270.1 GI:14425900
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
REFERENCE 1 (bases 1 to 396)
AUTHORS Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R., Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F., Goldman, G.H., Carvalho, A.F., Matsukuma, A., Baia, G.S., Simpson, D.H., Brunstein, A., de Oliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and Simpson, A.J.
TITLE Shotgun sequencing of the human transcriptome with ORF expressed sequence tags
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE 20202663
PUBMED 10737800
COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?cl=PM3&t2=PM3-MT0200-060101-004-g12&t3=2001-01-06&t4=1)
Seq primer: puc 18 forward
High quality sequence start: 20
High quality sequence stop: 396.
FEATURES
source
1..396
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="Adult"
/clone_lib="MT0200"
/note="Organ: marrow; Vector: puc18; Site: 1: Smal; Site 2: Smal; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the puc 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."
BASE COUNT 110 a 100 c 119 g 67 t
ORIGIN
Query Match 43.6%; Score 130.8; DB 12; Length 396;
Best Local Similarity 92.0%; Pred. No. 1.3e-21;
Matches 138; Conservative 0; Mismatches 12; Indels 0; Gaps 0;
QY 146 GCAGGAACGCTGAACCTGAGCGGAGACAGACCACTAGAGATGAGAGTCTGCTTT 205
DB 110 GGAGCTGACTGATACACTCCAGCGGAGACAGACCACTAGAGATGAGAGTCTGCTTT 169
QY 206 GCAGACCGAGATTGCCAATCTGCTGAAGGAGAGAGAACTAGAGTTCATCTTGGCAGC 265
DB 170 GCAGACCGAGATTGCCAATCTGCTGAAGGAGAGAGAACTAGAGTTCATCTTGGCAGC 229
QY 266 TCACCGACCTGCTGCAAGATCCCTGATTA 295
DB 230 TCACCGACCTGCTGCAAGATCCCTGATGA 259
RESULT 10
BM711302
LOCUS BM711302 519 bp mRNA linear EST 28-FEB-2002
DEFINITION UI-E-DX1-agu-k-20-0-UI-r1 UI-E-DX1 Homo sapiens cDNA clone
ACCESSION UI-E-DX1-agu-k-20-0-UI 5', mRNA sequence.
VERSION BM711302
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 16, 2003, 00:58:11 ; Search time 204 Seconds
(without alignments)
4806.183 Million cell updates/sec

Title: US-10-059-720-18

Perfect score: 300

Sequence: 1 CCATGGACTACAGGACGAC.....CAAGATCCCTGATTAGCTT 300

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2169961 seqs, 1634102185 residues

Total number of hits satisfying chosen parameters: 4339922

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications NA:*

- 1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
- 3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
- 4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
- 5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
- 6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
- 7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
- 8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
- 10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
- 11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
- 12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
- 13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
- 14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
- 15: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
- 16: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
- 17: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	300	100.0	300	14	US-10-059-720-18
2	267.2	89.1	281	14	Sequence 18, Appl
3	179.4	59.8	324	14	Sequence 16, Appl
4	160.6	53.5	266	14	Sequence 14, Appl
5	133.4	44.5	262	14	Sequence 10, Appl
6	130.8	43.6	426	11	US-10-059-720-10
7	130.8	43.6	2084	12	US-09-918-995-33976
8	130.8	43.6	2135	12	US-09-918-995-33976
9	130.2	43.4	273	14	Sequence 318, Appl
10	127.6	42.5	473	11	Sequence 8, Appl
11	127.4	42.5	451	11	Sequence 16, Appl
12	124.4	41.5	435	11	Sequence 34766, A
13	124.4	41.5	496	12	US-09-918-995-34766
14	124.4	41.5	600	12	Sequence 27268, A
15	124.4	41.5	1946	9	Sequence 1, Appl
16	124.4	41.5	2227	13	Sequence 25813, A
					Sequence 12113, A
					Sequence 3420, Ad
					Sequence 36, Appl

17	124.4	41.5	3565	12	US-09-873-319-649	Sequence 649, App
18	124.4	41.5	3565	12	US-09-960-706-993	Sequence 993, App
19	124.4	41.5	3565	12	US-09-873-367C-793	Sequence 793, App
20	124.4	41.5	6210	12	US-10-101-510-201	Sequence 201, App
21	123.4	41.1	642	9	US-09-864-761-20191	Sequence 20191, A
22	122.8	40.9	445	11	US-09-918-995-10656	Sequence 10656, A
23	121.4	40.5	296	14	US-10-059-720-32	Sequence 32, Appl
24	118	39.3	403	10	US-09-960-352-11134	Sequence 11134, A
25	116.4	38.8	399	10	US-09-960-352-7855	Sequence 7855, Ap
26	116.4	38.8	1176	13	US-10-002-600-39	Sequence 39, Appl
27	114.6	38.2	497	14	US-10-066-543-2435	Sequence 2435, Ap
28	112.8	37.9	429	11	US-09-918-995-34550	Sequence 34550, A
29	112.4	37.5	300	10	US-09-920-300A-436	Sequence 436, App
30	112.4	37.5	300	12	US-10-099-926-436	Sequence 436, App
31	112.4	37.5	300	13	US-10-033-528-436	Sequence 436, App
32	106.8	35.6	415	10	US-09-960-352-14906	Sequence 14906, A
33	105.2	35.1	3967	11	US-09-165-522-15	Sequence 15, Appl
34	104.4	34.8	120	10	US-09-912-787-23	Sequence 23, Appl
35	104.4	34.8	120	13	US-10-014-326-8	Sequence 8, Appl
36	94	31.3	918	10	US-09-912-787-5	Sequence 5, Appl
37	94	31.3	918	13	US-10-014-326-69	Sequence 69, Appl
38	94	31.3	925	10	US-09-912-787-13	Sequence 13, Appl
39	94	31.3	925	10	US-09-912-787-72	Sequence 72, Appl
40	94	31.3	925	13	US-10-014-326-58	Sequence 58, Appl
41	94	31.3	925	13	US-10-014-326-77	Sequence 77, Appl
42	94	31.3	928	10	US-09-912-787-9	Sequence 9, Appl
43	94	31.3	928	10	US-09-912-787-68	Sequence 68, Appl
44	94	31.3	928	10	US-09-912-787-74	Sequence 74, Appl
45	94	31.3	928	13	US-10-014-326-54	Sequence 54, Appl

ALIGNMENTS

RESULT 1

US-10-059-720-18
; Sequence 18, Application US/10059720
; Publication No. US20030027314A1
; GENERAL INFORMATION:

APPLICANT: VINSON, Charles R.
KRYLOV, Dmitry

TITLE OF INVENTION: EXTENSION OF A PROTEIN-PROTEIN
INTERACTION SURFACE TO INACTIVATE THE FUNCTION OF A
CELLULAR PROTEIN

NUMBER OF SEQUENCES: 64

CORRESPONDENCE ADDRESS:

ADDRESSEE: MORGAN & FINNEGAN, L.L.P.

STREET: 345 Park Avenue

CITY: New York

STATE: NY

COUNTRY: USA

ZIP: 10154-0053

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: MS WORD 97

SOFTWARE: ASCII

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/10/059,720

FILING DATE: 29-Jan-2002

PRIOR APPLICATION NUMBER:

APPLICATION NUMBER: 60/001,654

FILING DATE: 31-JUL-1995

APPLICATION NUMBER: 60/018,496

FILING DATE: 29-MAY-1996

APPLICATION NUMBER: 08/690,011

FILING DATE: 31-JULY-1996

ATTORNEY/AGENT INFORMATION:

NAME: Serunian, Leslie A.

REGISTRATION NUMBER: 35,353

REFERENCE/DOCKET NUMBER: 2026-4199US2

TELECOMMUNICATION INFORMATION:

TELEPHONE: (212) 758-4800

TELEFAX: (212)751-6849
INFORMATION FOR SEQ ID NO: 18:
SEQUENCE CHARACTERISTICS:
LENGTH: 300 base pairs
TYPE: nucleic acid
STRANDEDNESS: unknown
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
SEQUENCE DESCRIPTION: SEQ ID NO: 18:
US-10-059-720-18

Query Match 100.0%; Score 300; DB 14; Length 300;
Best Local Similarity 100.0%; Pred. No. 8.1e-80;
Matches 300; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CCATGGACTACAGGACGACGATGACGACGATATGGCTAGCATGCTGGGACGACAA 60
DB 1 CCATGGACTACAGGACGACGATGACGACGATATGGCTAGCATGCTGGGACGACAA 60
QY 61 TGGTTCGGATCCCTGACCTGCAACAGCTGCTGAGGAACTGGCCGCGGAAACGAGAGC 120
DB 61 TGGTTCGGATCCCTGACCTGCAACAGCTGCTGAGGAACTGGCCGCGGAAACGAGAGC 120
QY 121 TGGAAAGAGCGCCGAGAGAGCTGGAGCAGGAGAAACGCTGAACTCGAGCGGAGACGACC 180
DB 121 TGGAAAGAGCGCCGAGAGAGCTGGAGCAGGAGAAACGCTGAACTCGAGCGGAGACGACC 180
QY 181 AACTAGAGATGAGAGTCTCTTTGACGACGAGATGCGCACTCTCTGAGAGAGG 240
DB 181 AACTAGAGATGAGAGTCTCTTTGACGACGAGATGCGCACTCTCTGAGAGAGG 240
QY 241 AAAAAGTATGATTCATCTGCGAGCTCAGGACCTGCGTCAAGATCCCTGATTAGCTT 300
DB 241 AAAAAGTATGATTCATCTGCGAGCTCAGGACCTGCGTCAAGATCCCTGATTAGCTT 300

RESULT 2

US-10-059-720-16
Sequence 16, Application US/10059720
Publication No. US20030027314A1
GENERAL INFORMATION:
APPLICANT: VINSON, Charles R.
KEYLOV, Dmitry

TITLE OF INVENTION: EXTENSION OF A PROTEIN-PROTEIN
INTERACTION SURFACE TO INACTIVATE THE FUNCTION OF A
CELLULAR PROTEIN

NUMBER OF SEQUENCES: 64

CORRESPONDENCE ADDRESS:

ADDRESSEE: MORGAN & FINNEGAN, L.L.P.
STREET: 345 Park Avenue
CITY: New York
STATE: NY

COUNTRY: USA

ZIP: 10154-0053

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: MS WORD 97

SOFTWARE: ASCII

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/10/059,720

FILING DATE: 29-Jan-2002

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 60/001,654

FILING DATE: 31-JUL-1995

APPLICATION NUMBER: 60/018,496

FILING DATE: 29-MAY-1996

APPLICATION NUMBER: 08/690,011

FILING DATE: 31-JULY-1996

ATTORNEY/AGENT INFORMATION:

NAME: Serunian, Leslie A.

REGISTRATION NUMBER: 35,353

REFERENCE/DOCKET NUMBER: 2026-4199US2

TELECOMMUNICATION INFORMATION:

TELEPHONE: (212)758-4800

TELEFAX: (212)751-6849

INFORMATION FOR SEQ ID NO: 16:

SEQUENCE CHARACTERISTICS:

LENGTH: 281 base pairs

TYPE: nucleic acid

STRANDEDNESS: unknown

TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)

SEQUENCE DESCRIPTION: SEQ ID NO: 16:

US-10-059-720-16

Query Match 89.1%; Score 267.2; DB 14; Length 281;
Best Local Similarity 98.9%; Pred. No. 5e-70;
Matches 269; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 25 ACNACATATGCTAGCATGCTGTCACAGCAATATGGTGGGATCCCTGACCTGGAAC 84
DB 1 ATATACATATGCTAGCATGCTGTCACAGCAATATGGTGGGATCCCTGACCTGGAAC 60
QY 85 AACGTGCTGAGAACTGSCCGGTGAAACGAGAGCTGGAAAAAGAGCGCCGAGAGCTGG 144
DB 61 AACGTGCTGAGAACTGSCCGGTGAAACGAGAGCTGGAAAAAGAGCGCCGAGAGCTGG 120
QY 145 AGCAGAAACGCTGACTGCTGAGCGGAGACCACTAGAGATGAGAGATGAGAGCTGCTT 204
DB 121 ACAGAAACGCTGACTGCTGAGCGGAGACCACTAGAGATGAGAGATGAGAGCTGCTT 180
QY 205 TGCAGACGAGATTGCCAACCTGCTGAAGGAGAGAGGAAAAAATAGAGTTTATCTGSCAG 264
DB 181 TGCAGACGAGATTGCCAACCTGCTGAAGGAGAGAGGAAAAAATAGAGTTTATCTGSCAG 240
QY 265 CTCACGACCTGCTGAGAGTCCCTGATTAA 296
DB 241 CTCACGACCTGCTGAGAGTCCCTGATTAA 272

RESULT 3

US-10-059-720-14

Sequence 14, Application US/10059720

Publication No. US20030027314A1

GENERAL INFORMATION:

APPLICANT: VINSON, Charles R.

KEYLOV, Dmitry

TITLE OF INVENTION: EXTENSION OF A PROTEIN-PROTEIN

INTERACTION SURFACE TO INACTIVATE THE FUNCTION OF A

CELLULAR PROTEIN

NUMBER OF SEQUENCES: 64

CORRESPONDENCE ADDRESS:

ADDRESSEE: MORGAN & FINNEGAN, L.L.P.

STREET: 345 Park Avenue

CITY: New York

STATE: NY

COUNTRY: USA

ZIP: 10154-0053

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: MS WORD 97

SOFTWARE: ASCII

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/10/059,720

FILING DATE: 29-Jan-2002

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 60/001,654

FILING DATE: 31-JUL-1995

APPLICATION NUMBER: 60/018,496

FILING DATE: 29-MAY-1996

APPLICATION NUMBER: 08/690,011

FILING DATE: 31-JULY-1996

ATTORNEY/AGENT INFORMATION:

NAME: Serunian, Leslie A.

REGISTRATION NUMBER: 35,353
 REFERENCE/DOCKET NUMBER: 2026-4199US2
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (212)758-4800
 TELEFAX: (212)751-6849
 INFORMATION FOR SEQ ID NO: 14:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 324 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: unknown
 TOPOLOGY: linear
 MOLECULE TYPE: DNA (genomic)
 SEQUENCE DESCRIPTION: SEQ ID NO: 14:

US-10-059-720-14

Query Match 59.8%; Score 179.4; DB 14; Length 324;
 Best Local Similarity 77.8%; Pred. No. 8.9e-44;
 Matches 253; Conservative 0; Mismatches 46; Indels 26; Gaps 2;

QY 1 CCATGGACTACAAGGACGACGATGACAGCATATGGCTAGCATGACTGGTGACAGCAAA 60
 DB 1 CCATGGACTACAAGGACGACGATGACAGCATATGGCTAGCATGACTGGTGACAGCAAA 60
 QY 61 TGGGTCCGGATCTGACCTGCAACACGTGCTGAGGACTGGCCCGTGAACGAGAGC 120
 DB 61 TGGGTCCGGATCTGACCTGCAACACGTGCTGAGGACTGGCCCGTGAACGAGAGC 119
 QY 121 TGGAAAAGAGCGCCGAGAA-----GCTGGAGCAGGAAAC 155
 DB 120 CGAAGGGAAGGAATAGATGGCTGCACCAATGCCGACCGAGGAGGAGTCACT 179
 QY 156 GCTGAACCTGAGGCGGAGACAGCAACTAGAAATGAGAGTCTGCTTTGACACCGAG 215
 DB 180 GATACACTCCAGCGGAGACAGCAACTAGAAATGAGAGTCTGCTTTGACACCGAG 239
 QY 216 ATTGCCAACCTGCTGAAGGAGAGGAAAACTAGAGTTCACTCGGAGCTCACCGACCT 275
 DB 240 ATTGCCAACCTGCTGAAGGAGAGGAAAACTAGAGTTCACTCGGAGCTCACCGACCT 299
 QY 276 GCCTGCAAGATCCCTGATTAAAGCTT 300
 DB 300 GCCTGCAAGATCCCTGATTAAAGCTT 324

RESULT 4

US-10-059-720-10
 Sequence 10, Application US/10059720
 Publication No. US20030027314A1
 GENERAL INFORMATION:
 APPLICANT: VINSON, Charles R.
 TITLE OF INVENTION: EXTENSION OF A PROTEIN-PROTEIN
 INTERACTION SURFACE TO INACTIVATE THE FUNCTION OF A
 CELLULAR PROTEIN
 NUMBER OF SEQUENCES: 64
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: MORGAN & FINNEGAN, L.L.P.
 STREET: 345 Park Avenue
 CITY: New York
 STATE: NY
 COUNTRY: USA
 ZIP: 10154-0053
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: MS WORD 97
 SOFTWARE: ASCII
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/10/059,720
 FILING DATE: 29-Jan-2002
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 60/001,654
 FILING DATE: 31-JUL-1995

APPLICATION NUMBER: 60/018,496
 FILING DATE: 29-MAY-1996
 APPLICATION NUMBER: 08/690,011
 FILING DATE: 31-JULY-1996
 ATTORNEY/AGENT INFORMATION:
 NAME: Serunian, Leslie A.
 REGISTRATION NUMBER: 35,353
 REFERENCE/DOCKET NUMBER: 2026-4199US2
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (212)758-4800
 TELEFAX: (212)751-6849
 INFORMATION FOR SEQ ID NO: 10:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 266 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: unknown
 TOPOLOGY: linear
 MOLECULE TYPE: DNA (genomic)
 SEQUENCE DESCRIPTION: SEQ ID NO: 10:

US-10-059-720-10

Query Match 53.5%; Score 160.6; DB 14; Length 266;
 Best Local Similarity 97.6%; Pred. No. 3.5e-38;
 Matches 163; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 CCATGGACTACAAGGACGACGATGACAGCATATGGCTAGCATGACTGGTGACAGCAAA 60
 DB 1 CCATGGACTACAAGGACGACGATGACAGCATATGGCTAGCATGACTGGTGACAGCAAA 60
 QY 61 TGGGTCCGGATCTGACCTGCAACACGTGCTGAGGACTGGCCCGTGAACGAGAGC 120
 DB 61 TGGGTCCGGATCTGACCTGCAACACGTGCTGAGGACTGGCCCGTGAACGAGAGC 120
 QY 121 TGGAAAAGAGCGCCGAGAACTAGAGTTCACTCGGAGCTGAACTCGAG 167
 DB 121 TGGAAAAGAGCGCCGAGAACTAGAGTTCACTCGGAGCTGAACTCGAG 167

RESULT 5

US-10-059-720-12
 Sequence 12, Application US/10059720
 Publication No. US20030027314A1
 GENERAL INFORMATION:
 APPLICANT: VINSON, Charles R.
 TITLE OF INVENTION: EXTENSION OF A PROTEIN-PROTEIN
 INTERACTION SURFACE TO INACTIVATE THE FUNCTION OF A
 CELLULAR PROTEIN
 NUMBER OF SEQUENCES: 64
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: MORGAN & FINNEGAN, L.L.P.
 STREET: 345 Park Avenue
 CITY: New York
 STATE: NY
 COUNTRY: USA
 ZIP: 10154-0053
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: MS WORD 97
 SOFTWARE: ASCII
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/10/059,720
 FILING DATE: 29-Jan-2002
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 60/001,654
 FILING DATE: 31-JUL-1995
 APPLICATION NUMBER: 60/018,496
 FILING DATE: 29-MAY-1996
 APPLICATION NUMBER: 08/690,011
 FILING DATE: 31-JULY-1996
 ATTORNEY/AGENT INFORMATION:
 NAME: Serunian, Leslie A.

Db 607 GGAGTGTACTATACATCCAGGGGAGACAGACCACTAGAGATGAGAGTCTGCTTT 666
QY 206 GCAGACCGAGATTGCCAACCTGCTGAGGAGAGAGAACTAGATTCTCTGGCAGC 265
Db 667 GCAGACCGAGATTGCCAACCTGCTGAGGAGAGAGAACTAGATTCTCTGGCAGC 726
QY 266 TCACCGACCTGCTGCAAGATCCCTGATTA 295
Db 727 TCACCGACCTGCTGCAAGATCCCTGATTA 756

RESULT 9

US-10-070-676-16/c
; Sequence 16, Application US/10070676
; Publication No. US20030059788A1
; GENERAL INFORMATION:
; APPLICANT: Toque, Bruno
; APPLICANT: Bracco, Laurent
; APPLICANT: Schweighoffer, Fabien
; TITLE OF INVENTION: Genetic Markers of Toxicity, Preparation
; TITLE OF INVENTION: and Uses
; FILE REFERENCE: 50146/003002
; CURRENT APPLICATION NUMBER: US/10/070,676
; CURRENT FILING DATE: 2002-03-06
; PRIOR APPLICATION NUMBER: PCT/FR00/02503
; PRIOR FILING DATE: 2000-09-12
; PRIOR APPLICATION NUMBER: FR 99/11405
; PRIOR FILING DATE: 1999-09-13
; PRIOR APPLICATION NUMBER: 09/456,370
; PRIOR FILING DATE: 1999-12-08
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16
; LENGTH: 273
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-070-676-16

Query Match 43.4%; Score 130.2; DB 14; Length 273;
Best Local Similarity 97.8%; Pred. No. 4.3e-29;
Matches 132; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 161 ACTCGAGCGCAGACAGACCACTAGAGATGAGAGTCTGCTTGCAGACCGAGATTGC 220
Db 237 ACTCGAGCGCAGACAGACCACTAGAGATGAGAGTCTGCTTGCAGACCGAGATTGC 178
QY 221 CAACCTGCTGAGGAGAGAGAGAACTAGAGTTCATCTGCGAGCTCACCGACTGCGTG 280
Db 177 CAACCTGCTGAGGAGAGAGAGAACTAGAGTTCATCTGCGAGCTCACCGACTGCGTG 118
QY 281 CAAGATCCCTGATTA 295
Db 117 CAAGATCCCTGATTA 103

RESULT 10

US-09-918-995-34766
; Sequence 34766, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; FILE REFERENCE: 20411-756
; CURRENT APPLICATION NUMBER: US/09/918,995
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US/09/235,076
; PRIOR FILING DATE: 1999-01-20
; NUMBER OF SEQ ID NOS: 38054
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 34766
; LENGTH: 435

; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-918-995-34766

Query Match 42.5%; Score 127.6; DB 11; Length 435;
Best Local Similarity 90.7%; Pred. No. 2.9e-28;
Matches 136; Conservative 0; Mismatches 14; Indels 0; Gaps 0;
QY 146 GCAGGAAACGCTGAACTCGAGCGGAGACAGACCACTAGAGATGAGAGTCTGCTTT 205
Db 149 GGAGCTGACTGATACATCTCCAAAGCGGAGACCACTAGAGATGAGAGTCTGCTTT 208
QY 206 GCAGACCGAGATTGCCAACCTGCTGAGGAGAGAGAACTAGATTCTCTGGCAGC 265
Db 209 GCAGACCGAGATTGCCAACCTGCTGAGGAGAGAGAACTAGAGATGAGAGTCTGCGC 268
QY 266 TCACCGACCTGCTGCAAGATCCCTGATTA 295
Db 269 TCACCGACCTGCTGCAAGATCCCTGATTA 298

RESULT 11

US-09-918-995-27268
; Sequence 27268, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; FILE REFERENCE: 20411-756
; CURRENT APPLICATION NUMBER: US/09/918,995
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US/09/235,076
; PRIOR FILING DATE: 1999-01-20
; NUMBER OF SEQ ID NOS: 38054
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 27268
; LENGTH: 451
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(451)
; OTHER INFORMATION: n = A,T,C or G
US-09-918-995-27268

Query Match 42.5%; Score 127.6; DB 11; Length 451;
Best Local Similarity 90.7%; Pred. No. 3e-28;
Matches 136; Conservative 0; Mismatches 14; Indels 0; Gaps 0;
QY 146 GCAGGAAACGCTGAACTCGAGCGGAGACAGACCACTAGAGATGAGAGTCTGCTTT 205
Db 158 GGAGCTGACTGATACATCTCCAAAGCGGAGACCACTAGAGATGAGAGTCTGCTTT 217
QY 206 GCAGACCGAGATTGCCAACCTGCTGAGGAGAGAGAACTAGAGTTCATCTCTGGCAGC 265
Db 218 GCAGACCGAGATTGCCAACCTGCTGAGGAGAGAGAACTAGAGTTCATCTCTGGCAGC 277
QY 266 TCACCGACCTGCTGCAAGATCCCTGATTA 295
Db 278 TCACCGACCTGCTGCAAGATCCCTGATTA 307

RESULT 12

US-09-765-298A-1
; Sequence 1, Application US/09765298A
; Patent No. US20020137017A1
; GENERAL INFORMATION:
; APPLICANT: ARONHEIM, AMI
; TITLE OF INVENTION: METHOD FOR DETECTION PROTEIN-PROTEIN INTERACTIONS AND A KIT TH
; FILE REFERENCE: 108387.01
; CURRENT APPLICATION NUMBER: US/09/765,298A
; CURRENT FILING DATE: 2001-01-22

;; PRIOR APPLICATION NUMBER: IL 125456
;; PRIOR FILING DATE: 1998-07-22
;; PRIOR APPLICATION NUMBER: IL 128017
;; PRIOR FILING DATE: 1999-01-12
;; NUMBER OF SEQ ID NOS: 31
;; SOFTWARE: PatentIn version 3.1
;; SEQ ID NO 1
;; LENGTH: 435
;; TYPE: DNA
;; ORGANISM: Homo sapiens
US-09-765-298A-1

Query Match 41.5%; Score 124.4; DB 10; Length 435;
Best Local Similarity 99.2%; Pred. No. 2.6e-27;
Matches 125; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 170 GGAGCAGACCACTAGAGATGAGAAGTCTCTTTGCGAGCCGAGATTGCCAACCTGCT 229
|||||
Db 219 GGAGCAGACCACTAGAGATGAGAAGTCTCTTTGCGAGCCGAGATTGCCAACCTGCT 278
|||||
QY 230 GAAGGAGAGGAAACTAGAGTTTCATCTGCGAGCTCACCAGCTGCTGCAAGATCCC 289
|||||
Db 279 GAAGGAGAGGAAACTAGAGTTTCATCTGCGAGCTCACCAGCTGCTGCAAGATCCC 338
|||||
QY 290 TGATTA 295
|||||
Db 339 TGATGA 344

RESULT 13

US-10-029-386-25813
;; Sequence 25813, Application US/10029386
;; Publication No. US20030194704A1
;; GENERAL INFORMATION:
;; APPLICANT: Penn, Sharron G.
;; APPLICANT: Rank, David R.
;; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR G
;; TITLE OF INVENTION: EXPRESSION ANALYSIS TWO
;; FILE REFERENCE: AEOMICA-X-2
;; CURRENT APPLICATION NUMBER: US/10/029,386
;; CURRENT FILING DATE: 2001-12-20
;; NUMBER OF SEQ ID NOS: 34288
;; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
;; SEQ ID NO 25813
;; LENGTH: 496
;; TYPE: DNA
;; ORGANISM: Homo sapiens
;; FEATURE:
;; OTHER INFORMATION: MAP TO CHR14.3
;; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 19
;; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 27
;; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 2.9
;; OTHER INFORMATION: EST HUMAN HIT: B1908747.1, EVALUATE 0.00e+00
;; OTHER INFORMATION: SWISSPROT HIT: P01100, EVALUATE 8.00e-87
;; OTHER INFORMATION: NT HIT: AF11167.2, EVALUATE 0.00e+00

US-10-029-386-25813

Query Match 41.5%; Score 124.4; DB 12; Length 496;
Best Local Similarity 99.2%; Pred. No. 2.7e-27;
Matches 125; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 170 GGAGCAGACCACTAGAGATGAGAAGTCTCTTTGCGAGCCGAGATTGCCAACCTGCT 229
|||||
Db 1 GGAGCAGACCACTAGAGATGAGAAGTCTCTTTGCGAGCCGAGATTGCCAACCTGCT 60
|||||
QY 230 GAAGGAGAGGAAACTAGAGTTTCATCTGCGAGCTCACCAGCTGCTGCAAGATCCC 289
|||||
Db 61 GAAGGAGAGGAAACTAGAGTTTCATCTGCGAGCTCACCAGCTGCTGCAAGATCCC 120
|||||
QY 290 TGATTA 295
|||||
Db 121 TGATGA 126

RESULT 14

US-10-029-386-12113
;; Sequence 12113, Application US/10029386
;; Publication No. US20030194704A1
;; GENERAL INFORMATION:
;; APPLICANT: Penn, Sharron G.
;; APPLICANT: Rank, David R.
;; APPLICANT: Hanzel, David K.
;; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
;; TITLE OF INVENTION: EXPRESSION ANALYSIS TWO
;; FILE REFERENCE: AEOMICA-X-2
;; CURRENT APPLICATION NUMBER: US/10/029,386
;; CURRENT FILING DATE: 2001-12-20
;; NUMBER OF SEQ ID NOS: 34288
;; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
;; SEQ ID NO 12113
;; LENGTH: 600
;; TYPE: DNA
;; ORGANISM: Homo sapiens
;; FEATURE:
;; OTHER INFORMATION: MAP TO CHR14.3
;; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 19
;; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 27
;; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 2.9
;; OTHER INFORMATION: NT HIT: AF11167.2, EVALUATE 0.00e+00
;; OTHER INFORMATION: SWISSPROT HIT: P01100, EVALUATE 4.00e-84
;; OTHER INFORMATION: EST_HUMAN HIT: BF526890.1, EVALUATE 0.00e+00

US-10-029-386-12113

Query Match 41.5%; Score 124.4; DB 12; Length 600;
Best Local Similarity 99.2%; Pred. No. 2.9e-27;
Matches 125; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 170 GGAGCAGACCACTAGAGATGAGAAGTCTCTTTGCGAGCCGAGATTGCCAACCTGCT 229
|||||
Db 69 GGAGCAGACCACTAGAGATGAGAAGTCTCTTTGCGAGCCGAGATTGCCAACCTGCT 128
|||||
QY 230 GAAGGAGAGGAAACTAGAGTTTCATCTGCGAGCTCACCAGCTGCTGCAAGATCCC 289
|||||
Db 129 GAAGGAGAGGAAACTAGAGTTTCATCTGCGAGCTCACCAGCTGCTGCAAGATCCC 188
|||||
QY 290 TGATTA 295
|||||
Db 189 TGATGA 194

RESULT 15

US-09-864-761-3420
;; Sequence 3420, Application US/09864761
;; Patent No. US20020048763A1
;; GENERAL INFORMATION:
;; APPLICANT: Penn, Sharron G.
;; APPLICANT: Rank, David R.
;; APPLICANT: Hanzel, David K.
;; APPLICANT: Chen, Wensheng
;; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FO
;; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
;; FILE REFERENCE: Aeomica-X-1
;; CURRENT APPLICATION NUMBER: US/09/864,761
;; CURRENT FILING DATE: 2001-05-23
;; PRIOR APPLICATION NUMBER: US 60/180,312
;; PRIOR FILING DATE: 2000-02-04
;; PRIOR APPLICATION NUMBER: US 60/207,456
;; PRIOR FILING DATE: 2000-05-26
;; PRIOR APPLICATION NUMBER: US 09/632,366
;; PRIOR FILING DATE: 2000-08-03
;; PRIOR APPLICATION NUMBER: GB 24263.6
;; PRIOR FILING DATE: 2000-10-04
;; PRIOR APPLICATION NUMBER: US 60/236,359
;; PRIOR FILING DATE: 2000-09-27
;; PRIOR APPLICATION NUMBER: PCT/US01/00666

```

; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annonax Sequence Listing Engine vers. 1.1
; SEQ ID NO 3420
; LENGTH: 1946
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AF111167.2
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 6.2
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 5.4
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 2.9
; OTHER INFORMATION: EXPRESSED IN HEL100, SIGNAL = 5.2
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 2.8
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 4.5
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 2.1
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 36
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.8
; US-09-864-761-3420
Query Match 41.5%; Score 124.4; DB 9; Length 1946;
Best Local Similarity 99.2%; Pred. No. 4.1e-27;
Matches 125; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 170 GGAGACAGACCAACTAGAGATGAGAGTCTGCTTGGAGACCGAGATTGCCAACCTGCT 229
Db 313 GGAGACAGACCAACTAGAGATGAGAGTCTGCTTGGAGACCGAGATTGCCAACCTGCT 372
Qy 230 GAAGGAGAGGAGAAACCTAGAGTTTCATCTGCGAGCTCACCGACCTGCTGCAAGATCCC 289
Db 373 GAAGGAGAGGAGAAACCTAGAGTTTCATCTGCGAGCTCACCGAGCTGCTGCAAGATCCC 432
Qy 290 TCATTA 295
Db 433 TGATCA 438
Search completed: November 16, 2003, 01:36:23
Job time : 205 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 16, 2003, 00:10:40 ; Search time 50 Seconds
(without alignments)
2648.299 Million cell updates/sec

Title: US-10-059-720-18

Perfect score: 300
Sequence: 1 CCGTGGACTACAGGACGAC.....CAAGTCCCTGATTAGCTT 300

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA:*

- 1: /cgn2_6/ptodata/1/ina/5A COMB.seq.*
- 2: /cgn2_6/ptodata/1/ina/5B COMB.seq.*
- 3: /cgn2_6/ptodata/1/ina/6A COMB.seq.*
- 4: /cgn2_6/ptodata/1/ina/6B COMB.seq.*
- 5: /cgn2_6/ptodata/1/ina/PCTUS COMB.seq.*
- 6: /cgn2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	300	100.0	300	4	US-09-299-495F-18
2	294	98.0	294	2	US-08-690-011A-18
3	287.2	89.1	281	4	US-09-299-495F-16
4	285.6	88.5	281	2	US-08-690-011A-16
5	179.4	59.8	324	4	US-09-299-495F-14
6	173.4	57.8	318	2	US-08-690-011A-14
7	160.6	53.5	266	4	US-09-299-495F-10
8	157	52.3	264	2	US-08-690-011A-10
9	133.4	44.5	262	2	US-08-690-011A-12
10	133.4	44.5	262	4	US-09-299-495F-12
11	124.4	41.5	435	4	US-08-765-298A-1
12	124.4	41.5	3565	1	US-08-306-691B-44
13	124.4	41.5	3565	5	PCT-US93-06251-20
14	121.4	40.5	296	2	US-08-690-011A-32
15	121.4	40.5	296	4	US-09-299-495F-32
16	91.6	30.5	151	2	US-07-801-798-4
17	91.6	30.5	151	5	PCT-US92-10140-4
18	83.4	27.8	290	4	US-09-299-495F-8
19	81.4	27.1	288	2	US-08-690-011A-8
20	79.4	26.5	954	3	US-09-418-641-3
21	78.2	26.1	347	4	US-09-299-495F-28
22	76.6	25.5	347	2	US-08-690-011A-28
23	73.8	24.6	4144	1	US-08-218-686-1
24	73.8	24.6	4144	3	US-08-460-242-1
25	73.6	24.5	344	2	US-08-690-011A-24
26	73.6	24.5	344	4	US-09-299-495F-26
27	73.2	24.4	353	4	US-09-299-495F-20

28	72.6	24.2	180	2	US-08-256-790-9	Sequence 9, Appl
29	72.6	24.2	180	2	US-08-256-790-13	Sequence 13, Appl
30	72.2	24.1	267	4	US-09-059-625-5	Sequence 5, Appl
31	72.2	24.1	729	4	US-09-059-625-33	Sequence 33, Appl
32	72.2	24.1	729	4	US-09-059-625-44	Sequence 44, Appl
33	72.2	24.1	729	4	US-09-059-625-47	Sequence 47, Appl
34	71.6	23.9	341	4	US-09-299-495F-24	Sequence 24, Appl
35	71.2	23.7	462	3	US-08-863-813A-35	Sequence 35, Appl
36	71.2	23.7	462	3	US-08-676-318A-35	Sequence 35, Appl
37	70	23.3	341	2	US-08-690-011A-26	Sequence 26, Appl
38	69.6	23.2	345	2	US-08-690-011A-20	Sequence 20, Appl
39	55	18.3	7218	1	US-08-232-463-14	Sequence 14, Appl
40	53.8	17.9	324	4	US-09-299-495F-22	Sequence 22, Appl
41	50.6	16.9	324	2	US-08-690-011A-22	Sequence 22, Appl
42	49.6	16.5	456	2	US-08-557-309B-16	Sequence 16, Appl
43	49.6	16.5	456	3	US-08-834-304A-16	Sequence 16, Appl
44	49.6	16.5	456	3	US-08-993-674A-16	Sequence 16, Appl
45	49.6	16.5	456	4	US-09-256-976-16	Sequence 16, Appl

ALIGNMENTS

RESULT 1

US-09-299-495F-18

; Sequence 18, Application US/09299495F

; Patent No. 6361968

; GENERAL INFORMATION:

; APPLICANT: VINSON, Charles R.

; KRYLOV, Dmitry

; TITLE OF INVENTION: EXTENSION OF A PROTEIN-PROTEIN

; INTERACTION SURFACE TO INACTIVATE THE FUNCTION OF A

; CELLULAR PROTEIN

; NUMBER OF SEQUENCES: 64

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: MORGAN & FINNEGAN, L.L.P.

; STREET: 345 Park Avenue

; CITY: New York

; STATE: NY

; COUNTRY: USA

; ZIP: 10154-0053

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: MS WORD 97

; SOFTWARE: ASCII

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/299,495F

; FILING DATE: 26-Apr-1999

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 60/001,654

; FILING DATE: 31-JUL-1995

; APPLICATION NUMBER: 60/018,496

; FILING DATE: 29-MAY-1996

; APPLICATION NUMBER: 08/690,011

; FILING DATE: 31-JULY-1996

; ATTORNEY/AGENT INFORMATION:

; NAME: Serunian, Leslie A.

; REGISTRATION NUMBER: 35,353

; REFERENCE/DOCKET NUMBER: 2026-4199US2

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (212)758-4800

; TELEFAX: (212)751-6849

; INFORMATION FOR SEQ ID NO: 18:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 300 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: unknown

; TOPOLOGY: linear

; MOLECULE TYPE: DNA (genomic)

; SEQUENCE DESCRIPTION: SEQ ID NO: 18:

US-09-299-495F-18

Query Match 100.0%; Score 300; DB 4; Length 300;
Best Local Similarity 100.0%; Pred. No. 3.3e-73;
Matches 300; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 CCATGACTACAGGACGACGATGACACATATGGCTAGCATGACTGTGGACAGCAA 60
DB 1 CCATGACTACAGGACGACGATGACACATATGGCTAGCATGACTGTGGACAGCAA 60
QY 61 TGGGTCCGGATCTCGACCTGGAAACACGCTGAGGAACTGGCCCGTGAACGAGAGC 120
DB 61 TGGGTCCGGATCTCGACCTGGAAACACGCTGAGGAACTGGCCCGTGAACGAGAGC 120
QY 121 TGGAAAGAGGCGGAGAGCTGGAGCAGGAAACCGCTGAACCTGGAGGCGGAGAGC 180
DB 121 TGGAAAGAGGCGGAGAGCTGGAGCAGGAAACCGCTGAACCTGGAGGCGGAGAGC 180
QY 181 AACTAGAGATGAGAGTCTGCTTTCAGACCGAGATTGCCAACCCTGCTGAAGGAGG 240
DB 181 AACTAGAGATGAGAGTCTGCTTTCAGACCGAGATTGCCAACCCTGCTGAAGGAGG 240
QY 241 AAAAAGTAGAGTTTCATCTGGCAGCTCAGCGAGCTCCCTGCAAGATCCCTGATTAGCTT 300
DB 241 AAAAAGTAGAGTTTCATCTGGCAGCTCAGCGAGCTCCCTGCAAGATCCCTGATTAGCTT 300
```

RESULT 2

US-08-690-011A-18
; Sequence 18, Application US/08690011A
; Patent No. 5942433

GENERAL INFORMATION:

APPLICANT: VINSON, Charles R.

APPLICANT: KRYLOV, Dmitry

TITLE OF INVENTION: EXTENSION OF A PROTEIN-PROTEIN

TITLE OF INVENTION: INTERACTION SURFACE TO INACTIVATE THE FUNCTION OF A

TITLE OF INVENTION: CELLULAR PROTEIN

NUMBER OF SEQUENCES: 60

CORRESPONDENCE ADDRESS:

ADDRESSEE: MORGAN & FINNEGAN, L.L.P.

STREET: 345 Park Avenue

CITY: New York

STATE: NY

COUNTRY: USA

ZIP: 10154-0053

COMPUTER READABLE FORM: disk

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent In Release #1.0, Version #1.30B

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/690,011A

FILING DATE: 31-JUL-1996

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 60/001,654

FILING DATE: 31-JUL-1995

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 60/018,496

FILING DATE: 29-MAY-1996

ATTORNEY/AGENT INFORMATION:

NAME: Serunian, Leslie A.

REGISTRATION NUMBER: 35,353

REFERENCE/DOCKET NUMBER: 2026-4199US1

TELEPHONE: (212)751-6849

TELEFAX: (212)751-6849

INFORMATION FOR SEQ ID NO: 18:

SEQUENCE CHARACTERISTICS:

LENGTH: 294 base pairs

TYPE: nucleic acid

STRANDEDNESS: unknown

TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)

US-08-690-011A-18

Query Match 98.0%; Score 294; DB 2; Length 294;
Best Local Similarity 100.0%; Pred. No. 1.4e-71;
Matches 294; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 3 ATGGACTACAGGACGACGATGACACATATGGCTAGCATGACTGTGGACAGCAATG 62
DB 1 ATGGACTACAGGACGACGATGACACATATGGCTAGCATGACTGTGGACAGCAATG 60
QY 63 GTCGGGATCTCGACCTGGAAACACGCTGAGGAACTGGCCCGTGAACGAGAGCTG 122
DB 61 GTCGGGATCTCGACCTGGAAACACGCTGAGGAACTGGCCCGTGAACGAGAGCTG 120
QY 123 GAAAGAGGCGGAGAGCTGGAGCAGGAAACCGCTGAACCTGGAGGCGGAGAGCA 182
DB 121 GAAAGAGGCGGAGAGCTGGAGCAGGAAACCGCTGAACCTGGAGGCGGAGAGCA 180
QY 183 CTAGAGATGAGAGTCTGCTTTCAGACCGAGATTGCCAACCCTGCTGAAGGAGG 242
DB 181 CTAGAGATGAGAGTCTGCTTTCAGACCGAGATTGCCAACCCTGCTGAAGGAGG 240
QY 243 AAAGTAGAGTTTCATCTGGCAGCTCAGCGAGCTCCCTGCAAGATCCCTGATTAA 296
DB 241 AAAGTAGAGTTTCATCTGGCAGCTCAGCGAGCTCCCTGCAAGATCCCTGATTAA 294
```

RESULT 3

US-09-299-495F-16
; Sequence 16, Application US/09299495F
; Patent No. 6361668

GENERAL INFORMATION:

APPLICANT: VINSON, Charles R.

APPLICANT: KRYLOV, Dmitry

TITLE OF INVENTION: EXTENSION OF A PROTEIN-PROTEIN

TITLE OF INVENTION: INTERACTION SURFACE TO INACTIVATE THE FUNCTION OF A

TITLE OF INVENTION: CELLULAR PROTEIN

NUMBER OF SEQUENCES: 64

CORRESPONDENCE ADDRESS:

ADDRESSEE: MORGAN & FINNEGAN, L.L.P.

STREET: 345 Park Avenue

CITY: New York

STATE: NY

COUNTRY: USA

ZIP: 10154-0053

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: MS WORD 97

SOFTWARE: ASCII

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/299,495F

FILING DATE: 26-Apr-1999

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 60/001,654

FILING DATE: 31-JUL-1995

APPLICATION NUMBER: 60/018,496

FILING DATE: 29-MAY-1996

APPLICATION NUMBER: 08/690,011

FILING DATE: 31-JULY-1996

ATTORNEY/AGENT INFORMATION:

NAME: Serunian, Leslie A.

REGISTRATION NUMBER: 35,353

REFERENCE/DOCKET NUMBER: 2026-4199US2

TELEPHONE: (212)751-6849

TELEFAX: (212)751-6849

INFORMATION FOR SEQ ID NO: 16:

SEQUENCE CHARACTERISTICS:

LENGTH: 281 base pairs

TYPE: nucleic acid

STRANDEDNESS: unknown

TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)

SEQUENCE DESCRIPTION: SEQ ID NO: 16:

US-09-299-495F-16

Query Match 89.1%; Score 267.2; DB 4; Length 281;
Best Local Similarity 98.9%; Pred. No. 3e-64;
Matches 269; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 25 ACAAGCATATGGCTAGCATGACTGGTGGACAGCAAAATGGTGGGATCTCGACCTGGAAC 84
DB 1 ATATACATATGGCTAGCATGACTGGTGGACAGCAAAATGGTGGGATCTCGACCTGGAAC 60
QY 85 AACGTCTGAGGAACCTGGCCCGTGAACACGAAGAGCTGGAAAGAGGCGCGAAGAGCTGG 144
DB 61 AACGTCTGAGGAACCTGGCCCGTGAACACGAAGAGCTGGAAAGAGGCGCGAAGAGCTGG 120
QY 145 ACAGAGAAACGCTGAACTCGAGCGGAGACAGCAACCTGAGAGATGAGAGTCTGCTT 204
DB 121 ACAGAGAAACGCTGAACTCGAGCGGAGACAGCAACCTGAGAGATGAGAGTCTGCTT 180
QY 205 TGCAGACCGAGATTGCAACCTGCTGAAGGAGAGGAAACCTAGAGTTTCATCTGGCAG 264
DB 181 TGCAGACCGAGATTGCAACCTGCTGAAGGAGAGGAAACCTAGAGTTTCATCTGGCAG 240
QY 265 CTCACCGACCTGCTGCAAGATCCCTGATTAA 296
DB 241 CTCACCGACCTGCTGCAAGATCCCTGATTAA 272

RESULT 4

US-08-690-011A-16
; Sequence 16, Application US/08690011A
; Patent No. 5942433

GENERAL INFORMATION:

APPLICANT: VINSON, Charles R.

APPLICANT: KRYLOV, Dmitry

TITLE OF INVENTION: EXTENSION OF A PROTEIN-PROTEIN

TITLE OF INVENTION: INTERACTION SURFACE TO INACTIVATE THE FUNCTION OF A

TITLE OF INVENTION: CELLULAR PROTEIN

NUMBER OF SEQUENCES: 60

CORRESPONDENCE ADDRESS:

ADDRESSEE: MORGAN & FINNEGAN, L.L.P.

STREET: 345 Park Avenue

CITY: New York

STATE: NY

COUNTRY: USA

ZIP: 10154-0053

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent In Release #1.0, Version #1.30B

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/690,011A

FILING DATE: 31-JUL-1996

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 60/001,654

FILING DATE: 31-JUL-1995

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 60/018,496

FILING DATE: 29-MAY-1996

ATTORNEY/AGENT INFORMATION:

NAME: Serunian, Leslie A.

REGISTRATION NUMBER: 35,353

REFERENCE/DOCKET NUMBER: 2026-4199US1

TELECOMMUNICATION INFORMATION:

TELEPHONE: (212) 751-6840

TELEFAX: (212) 751-6849

INFORMATION FOR SEQ ID NO: 16:

SEQUENCE CHARACTERISTICS:

LENGTH: 281 base pairs

TYPE: nucleic acid

STRANDEDNESS: unknown

TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)

US-08-690-011A-16

Query Match 88.5%; Score 265.6; DB 2; Length 281;
Best Local Similarity 98.5%; Pred. No. 8.1e-64;
Matches 268; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 25 ACAAGCATATGGCTAGCATGACTGGTGGACAGCAAAATGGTGGGATCTCGACCTGGAAC 84
DB 1 ATATACATATGGCTAGCATGACTGGTGGACAGCAAAATGGTGGGATCTCGACCTGGAAC 60
QY 85 AACGTCTGAGGAACCTGGCCCGTGAACACGAAGAGCTGGAAAGAGGCGCGAAGAGCTGG 144
DB 61 AACGTCTGAGGAACCTGGCCCGTGAACACGAAGAGCTGGAAAGAGGCGCGAAGAGCTGG 120
QY 145 ACAGAGAAACGCTGAACTCGAGCGGAGACAGCAACCTAGAGATGAGAGTCTGCTT 204
DB 121 ACAGAGAAACGCTGAACTCGAGCGGAGACAGCAACCTAGAGATGAGAGTCTGCTT 180
QY 205 TGCAGACCGAGATTGCAACCTGCTGAAGGAGAGGAAACCTAGAGTTTCATCTGGCAG 264
DB 181 TGCAGACCGAGATTGCAACCTGCTGAAGGAGAGGAAACCTAGAGTTTCATCTGGCAG 240
QY 265 CTCACCGACCTGCTGCAAGATCCCTGATTAA 296
DB 241 CTCACCGACCTGCTGCAAGATCCCTGATTAA 272

RESULT 5

US-09-299-495F-14
; Sequence 14, Application US/09299495F
; Patent No. 6361968

GENERAL INFORMATION:

APPLICANT: VINSON, Charles R.

APPLICANT: KRYLOV, Dmitry

TITLE OF INVENTION: EXTENSION OF A PROTEIN-PROTEIN

TITLE OF INVENTION: INTERACTION SURFACE TO INACTIVATE THE FUNCTION OF A

TITLE OF INVENTION: CELLULAR PROTEIN

NUMBER OF SEQUENCES: 64

CORRESPONDENCE ADDRESS:

ADDRESSEE: MORGAN & FINNEGAN, L.L.P.

STREET: 345 Park Avenue

CITY: New York

STATE: NY

COUNTRY: USA

ZIP: 10154-0053

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: MS WORD 97

SOFTWARE: ASCII

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/299,495F

FILING DATE: 26-Apr-1999

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 60/001,654

FILING DATE: 31-JUL-1995

APPLICATION NUMBER: 60/018,496

FILING DATE: 29-MAY-1996

APPLICATION NUMBER: 08/690,011

FILING DATE: 31-JULY-1996

ATTORNEY/AGENT INFORMATION:

NAME: Serunian, Leslie A.

REGISTRATION NUMBER: 35,353

REFERENCE/DOCKET NUMBER: 2026-4199US2

TELECOMMUNICATION INFORMATION:

TELEPHONE: (212) 751-6840

TELEFAX: (212) 751-6849

INFORMATION FOR SEQ ID NO: 14:

SEQUENCE CHARACTERISTICS:

LENGTH: 324 base pairs

TYPE: nucleic acid

STRANDEDNESS: unknown

TOPOLOGY: linear

```
;
; MOLECULE TYPE: DNA (genomic)
; SEQUENCE DESCRIPTION: SEQ ID NO: 14:
US-09-299-495F-14

Query Match          59.8%; Score 179.4; DB 4; Length 324;
Best Local Similarity 77.4%; Pred. No. 3.1e-40;
Matches 253; Conservative 0; Mismatches 46; Indels 26; Gaps 2;

QY 1 CCATGGACTACAGGACGACGATGACAAAGCATATGGCTAGCATGCTGGGACGACAAA 60
   |||||
Db 1 CCATGGACTACAAAGGACGACGATGACAAAGCATATGGCTAGCATGCTGGGACGACAAA 60
   |||||

QY 61 TGGGTCGGGATCCTCACTGGAAACAACTGCTGAGAACTGCGCCGTTGAAACGAAAGC 120
   |||||
Db 61 TGGGTCGGGATCCTCACTGGAAACAACTGCTGAGAACTGCGCCGTTGAAACGAAAGC 120
   |||||

QY 121 TGGAAAAAGAGCCGGAAGA-----GCTGGAGACGAGAAAC 155
   |||||
Db 120 CGAAGGGAAGGATAGATGGCTGACGCCAAATGCCCAACCGGAGGAGGAGCTGACT 179
   |||||

QY 156 GCTGAACTCGAGCGGAGACAGACCACTAGAAAGATGAGAACTGCTTTTCAGACCGAG 215
   |||||
Db 180 GATACATCCAGCGGAGACAGACCACTAGAAAGATGAGAACTGCTTTTCAGACCGAG 239
   |||||

QY 216 ATTGCCAACTCTCTGAAGGAGAGGAAAAAAGTTCATCTCTGAGCTCAACCGACCT 275
   |||||
Db 240 ATTGCCAACTCTCTGAAGGAGAGGAAAAAAGTTCATCTCTGAGCTCAACCGACCT 299
   |||||

QY 276 GCTGCAAGATCCCTGATTAGCTT 300
   |||||
Db 300 GCTGCAAGATCCCTGATTAGCTT 324
   |||||

RESULT 6
US-08-690-011A-14
; Sequence 14, Application US/08690011A
; Patent No. 5942433
; GENERAL INFORMATION:
; APPLICANT: VINSON, Charles R.
; APPLICANT: KRYLOV, Dmitry
; TITLE OF INVENTION: EXTENSION OF A PROTEIN-PROTEIN
; TITLE OF INVENTION: INTERACTION SURFACE TO INACTIVATE THE FUNCTION OF A
; TITLE OF INVENTION: CELLULAR PROTEIN
; NUMBER OF SEQUENCES: 60
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORGAN & FINNEGAN, L.L.P.
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10154-0053
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30B
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/690,011A
; FILING DATE: 31-JUL-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/001,654
; FILING DATE: 31-JUL-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/018,496
; FILING DATE: 29-MAY-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Serunian, Leslie A.
; REGISTRATION NUMBER: 35,353
; REFERENCE/DOCKET NUMBER: 2026-4199US1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 751-6840
; TELEFAX: (212) 751-6849
; INFORMATION FOR SEQ ID NO: 14:

; SEQUENCE CHARACTERISTICS:
; LENGTH: 318 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: unknown
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-690-011A-14

Query Match          57.8%; Score 173.4; DB 2; Length 318;
Best Local Similarity 77.4%; Pred. No. 1.4e-38;
Matches 247; Conservative 0; Mismatches 46; Indels 26; Gaps 2;

QY 3 ATGGAATCAAGGACGACGATGACAAAGCATATGGCTAGCATGCTGGGACGACAAAATG 62
   |||||
Db 1 ATGGAATCAAGGACGACGATGACAAAGCATATGGCTAGCATGCTGGGACGACAAAATG 60
   |||||

QY 63 GGTGCGGATCTGACCTGGAAACAACTGCTGAGAACTGCGCCGTTGAAACGAAAGCTG 122
   |||||
Db 61 GGTGCGGATCTGACCTGGAAACAACTGCTGAGAACTGCGCCGTTGAAACGAAAGCTG 119
   |||||

QY 123 GAAAAAGAGCCGGAAGA-----GCTGGAGACGAGAAAC 157
   |||||
Db 120 AAGGAAAGATAGATGGCTGAGCCAAATGCCCAACCGGAGGAGGAGCTGACTGA 179
   |||||

QY 158 TGAATCTGAGCGGAGACAGACCACTAGAAAGATGAGAACTGCTTTTCAGACCGAGAT 217
   |||||
Db 180 TACATCTCAAGCGGAGACAGACCACTAGAAAGATGAGAACTGCTTTTCAGACCGAGAT 239
   |||||

QY 218 TGCACACTGCTGAGGAGAGGAAAAAAGTTCATCTCTGAGCTCAACCGACCTGC 277
   |||||
Db 240 TGCACACTGCTGAGGAGAGGAAAAAAGTTCATCTCTGAGCTCAACCGACCTGC 299
   |||||

QY 278 CTGCAAGATCCCTGATTAA 296
   |||||
Db 300 CTGCAAGATCCCTGATTAA 318
   |||||

RESULT 7
US-09-299-495F-10
; Sequence 10, Application US/09299495F
; Patent No. 6361968
; GENERAL INFORMATION:
; APPLICANT: VINSON, Charles R.
; APPLICANT: KRYLOV, Dmitry
; TITLE OF INVENTION: EXTENSION OF A PROTEIN-PROTEIN
; TITLE OF INVENTION: INTERACTION SURFACE TO INACTIVATE THE FUNCTION OF A
; TITLE OF INVENTION: CELLULAR PROTEIN
; NUMBER OF SEQUENCES: 64
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORGAN & FINNEGAN, L.L.P.
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10154-0053
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: MS WORD 97
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/299,495F
; FILING DATE: 26-APR-1999
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/001,654
; FILING DATE: 31-JUL-1995
; APPLICATION NUMBER: 60/018,496
; FILING DATE: 29-MAY-1996
; APPLICATION NUMBER: 08/690,011
; FILING DATE: 31-JULY-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Serunian, Leslie A.
; REGISTRATION NUMBER: 35,353
```

REFERENCE/DOCKET NUMBER: 2026-4199US2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212)758-4800
TELEFAX: (212)751-6849
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 266 base pairs
TYPE: nucleic acid
STRANDEDNESS: unknown
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
SEQUENCE DESCRIPTION: SEQ ID NO: 10:
US-09-299-495F-10

Query Match 53.5%; Score 160.6; DB 4; Length 266;
Best Local Similarity 97.6%; Pred. No. 4e-35;
Matches 163; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 1 CCATGACTACAAAGGACGACGATGACACCATATGCTAGCATGCTGGTGGACAGCAA 60
Db 1 CCATGACTACAAAGGACGACGATGACACCATATGCTAGCATGCTGGTGGACAGCAA 60
QY 61 TGGGTGGGATCTCTGACCTGGACAACTGCTGAGGAATGGCCCGTGAAAAACGAAGAGC 120
Db 61 TGGGTGGGATCTCTGACCTGGACAACTGCTGAGGAATGGCCCGTGAAAAACGAAGAGC 120
QY 121 TGGAAAAAGAGCCGAGAGCTGGACGACGAAACGCTGAACCTCGAG 167
Db 121 TGGAAAAAGAGCCGAGAGCTGGACGACGAAACGCTGGCAGAACTCGAG 167

RESULT 8

US-08-690-011A-10
Sequence 10, Application US/086900011A
Patent No. 5942433
GENERAL INFORMATION:
APPLICANT: VINSON, Charles R.
APPLICANT: KRYLOV, Dmitry
TITLE OF INVENTION: EXTENSION OF A PROTEIN-PROTEIN
TITLE OF INVENTION: INTERACTION SURFACE TO INACTIVATE THE FUNCTION OF A
TITLE OF INVENTION: CELLULAR PROTEIN
NUMBER OF SEQUENCES: 60
CORRESPONDENCE ADDRESS:
ADDRESSEE: MORGAN & FINNEGAN, L.L.P.
STREET: 345 Park Avenue
CITY: New York
STATE: NY
COUNTRY: USA
ZIP: 10154-0053
COMPUTER READABLE FORM:
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30B
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/690,011A
FILING DATE: 31-JUL-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/001,654
FILING DATE: 31-JUL-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/018,496
FILING DATE: 29-MAY-1996
ATTORNEY/AGENT INFORMATION:
NAME: Serunian, Leslie A.
REGISTRATION NUMBER: 35,353
REFERENCE/DOCKET NUMBER: 2026-4199US1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212)758-4800
TELEFAX: (212)751-6849
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 264 base pairs

TYPE: nucleic acid
STRANDEDNESS: unknown
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-690-011A-10
Query Match 52.3%; Score 157; DB 2; Length 264;
Best Local Similarity 97.0%; Pred. No. 3.9e-34;
Matches 160; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
QY 3 ATGGACTACAAAGGACGACGATGACACCATATGCTAGCATGCTGGTGGACAGCAAATG 62
Db 1 ATGGACTACAAAGGACGACGATGACACCATATGCTAGCATGCTGGTGGACAGCAAATG 60
QY 63 GGTGGGATCTCTGACCTGGAACTGCTGAGGAATGGCCCGTGAAAAACGAAGAGCTG 122
Db 61 GGTGGGATCTCTGACCTGGAACTGCTGAGGAATGGCCCGTGAAAAACGAAGAGCTG 120
QY 123 GAAAGAGGCGGAGAGCTGGACGACGAAACGCTGAACCTCGAG 167
Db 121 GAAAGAGGCGGAGAGCTGGACGACGAAACGCTGGCAGAACTCGAG 165

RESULT 9

US-08-690-011A-12
Sequence 12, Application US/086900011A
Patent No. 5942433
GENERAL INFORMATION:
APPLICANT: VINSON, Charles R.
APPLICANT: KRYLOV, Dmitry
TITLE OF INVENTION: EXTENSION OF A PROTEIN-PROTEIN
TITLE OF INVENTION: INTERACTION SURFACE TO INACTIVATE THE FUNCTION OF A
TITLE OF INVENTION: CELLULAR PROTEIN
NUMBER OF SEQUENCES: 60
CORRESPONDENCE ADDRESS:
ADDRESSEE: MORGAN & FINNEGAN, L.L.P.
STREET: 345 Park Avenue
CITY: New York
STATE: NY
COUNTRY: USA
ZIP: 10154-0053
COMPUTER READABLE FORM:
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30B
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/690,011A
FILING DATE: 31-JUL-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/001,654
FILING DATE: 31-JUL-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/018,496
FILING DATE: 29-MAY-1996
ATTORNEY/AGENT INFORMATION:
NAME: Serunian, Leslie A.
REGISTRATION NUMBER: 35,353
REFERENCE/DOCKET NUMBER: 2026-4199US1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212)758-4800
TELEFAX: (212)751-6849
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 262 base pairs
TYPE: nucleic acid
STRANDEDNESS: unknown
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-690-011A-12
Query Match 44.5%; Score 133.4; DB 2; Length 262;
Best Local Similarity 92.7%; Pred. No. 1.1e-27;

QY	206	GCAGACCGAGATTGCCACCTCTCTGACGAGAGCGAAAAA	CTAGAGTT	CATCTCTGCAGC	265
Db	163	GCAGACCGAGATTGCCACCTCTCTGACGAGAGCGAAAAA	CTAGAGTT	CATCTCTGCAGC	222
QY	266	TCACCGACCTGCCTGCAAGATCCCTGATTAA			296
Db	223	TCACCGACCTGCCTGCAAGATCCCTGATTAA			253

RESULT 11
US-09-765-298A-1
; Sequence 1, Application US/09765298A
Patent No. 6582927

```

/ AFFILIATION: ARROWHEAD, AND
/ TITLE OF INVENTION: METHOD FOR DETECTION PROTEIN-PROTEIN INTERACTIONS AND A
/ FILE REFERENCE: 108387.01
/ CURRENT APPLICATION NUMBER: US/09/765,298A
/ CURRENT FILING DATE: 2001-01-22
/ PRIOR APPLICATION NUMBER: IL 125456
/ PRIOR FILING DATE: 1998-07-22
/ PRIOR APPLICATION NUMBER: IL 128017
/ PRIOR FILING DATE: 1999-01-12
/ NUMBER OF SEQ ID NOS: 31
/ SOFTWARE: PatentIn version 3.1
/ SEQ ID NO 1
/ LENGTH: 435
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ US-09-765-298A-1

Query Match          41.5%; Score 124.4; DB 4; Length 435;
Best local similarity 99.2%; Pred. No. 3.7e-25;
Matches 125; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      170  GGAGACAGACCAACTAGATGAGATGAGAGTCTGCTTTGCAGAGCCGAGATTGCCAACCTGCT 229
      |||
Db       219  GGAGACAGACCAACTAGAGATGAGAGTCTGCTTTGCAGAGCCGAGATTGCCAACCTGCT 278
      |||

QY      230  GAAGGAGAGAGGAAAACTAGATTCTACTCTGCGAGCTCACCGACCTGCGCTGCAAGATCCC 289
      |||
Db       279  GAAGGAGAGAGGAAAACTAGATTCTACTCTGCGAGCTCACCGACCTGCGCTGCAAGATCCC 338
      |||

QY      290  TGATTA 295
      |||
Db       339  TGATGA 344

```

```

RESULT 12
US-08-306-691B-44
; Sequence 44, Application US/08306691B
; Patent No. 5734039
; GENERAL INFORMATION:
; APPLICANT: Calabretta, Bruno
; APPLICANT: Skorski, Tomasz
; TITLE OF INVENTION: ANTISENSE
; TITLE OF INVENTION: ANTISENSE
; NUMBER OF INVENTIONS: OLIGONUCLEOTIDES TARGETING COOPERATING ONCOGENES
; NUMBER OF SEQUENCES: 55
; CORRESPONDENCE ADDRESS:
; ADDRESSES: Seigel, Gonda, Lavorgna & Monaco, P.C.
; STREET: Two Penn Center, Suite 1800
; CITY: Philadelphia
; STATE: Pennsylvania
; COUNTRY: U.S.A.
; ZIP: 19102
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 720 Kb
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: WordPerfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/306,691B
; FILING DATE: September 15, 1994

```



```
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Monaco, Daniel A.
; REGISTRATION NUMBER: 30,480
; REFERENCE/DOCKET NUMBER: 8321-8
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (215) 568-8383
; TELEFAX: (215) 568-5549
; TELEX: No. 5734039e
; INFORMATION FOR SEQ ID NO: 44:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3565 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
US-08-306-691B-44

Query Match 41.5%; Score 124.4; DB 1; Length 3565;
Best Local Similarity 99.2%; Pred. No. 7.4e-25;
Matches 125; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 170 GGAGACAGACCAACTAGAGATGAGAACTCTGCTTTGCAGACCGAGATTGCCAACCTGCT 229
DB 2087 GGAGACAGACCAACTAGAGATGAGAACTCTGCTTTGCAGACCGAGATTGCCAACCTGCT 2146
QY 230 GAAGGAGAGAGAAAACCTAGAGTTTCATCTCTGGCAGCTCAGCGACTGCTGCTCAAGATCCC 289
DB 2147 GAAGGAGAGAGAAAACCTAGAGTTTCATCTCTGGCAGCTCAGCGACTGCTGCTCAAGATCCC 2206
QY 290 TGATTA 295
DB 2207 TGATGA 2212

RESULT 13
PCT-US93-06251-20
; Sequence 20, Application PC/TUS9306251
; GENERAL INFORMATION:
; APPLICANT: Wickstrom, Eric and Rife, Jason P.
; TITLE OF INVENTION: Trivalent Synthesis of Oligonucleotides Containing
; TITLE OF INVENTION: Stereospecific Alkylphosphonates and Arylphosphonates
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SCULLY, SCOTT, MURPHY & PRESSER
; STREET: 400 Garden City Plaza
; CITY: Garden City
; STATE: NY
; COUNTRY: USA
; ZIP: 11530
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/06251
; FILING DATE: 19930630
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: DiGiglio, Frank S.
; REGISTRATION NUMBER: 31,346
; REFERENCE/DOCKET NUMBER: 8586
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 516-742-4343
; TELEFAX: 516-742-4366
; TELEX: 230 901 SANS UR
; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3565 base pairs
; TYPE: nucleic acid
```

```
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
PCT-US93-06251-20

Query Match 41.5%; Score 124.4; DB 5; Length 3565;
Best Local Similarity 99.2%; Pred. No. 7.4e-25;
Matches 125; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 170 GGAGACAGACCAACTAGAGATGAGAACTCTGCTTTGCAGACCGAGATTGCCAACCTGCT 229
DB 2087 GGAGACAGACCAACTAGAGATGAGAACTCTGCTTTGCAGACCGAGATTGCCAACCTGCT 2146
QY 230 GAAGGAGAGAGAAAACCTAGAGTTTCATCTCTGGCAGCTCAGCGACTGCTGCTCAAGATCCC 289
DB 2147 GAAGGAGAGAGAAAACCTAGAGTTTCATCTCTGGCAGCTCAGCGACTGCTGCTCAAGATCCC 2206
QY 290 TGATTA 295
DB 2207 TGATGA 2212

RESULT 14
US-08-690-011A-32
; Sequence 32, Application US/086900011A
; Patent No. 5942433
; GENERAL INFORMATION:
; APPLICANT: VINSON, Charles R.
; APPLICANT: KRYLOV, Dmitry
; TITLE OF INVENTION: EXTENSION OF A PROTEIN-PROTEIN
; TITLE OF INVENTION: INTERACTION SURFACE TO INACTIVATE THE FUNCTION OF A
; TITLE OF INVENTION: CELLULAR PROTEIN
; NUMBER OF SEQUENCES: 60
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORGAN & FINNEGAN, L.L.P.
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10154-0053
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30B
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/690,011A
; FILING DATE: 31-JUL-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/001,654
; FILING DATE: 31-JUL-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/018,496
; FILING DATE: 29-MAY-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Serunian, Leslie A.
; REGISTRATION NUMBER: 35,353
; REFERENCE/DOCKET NUMBER: 2026-4199US1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 758-4800
; TELEFAX: (212) 751-6849
; INFORMATION FOR SEQ ID NO: 32:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 296 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: unknown
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-690-011A-32

Query Match 40.5%; Score 121.4; DB 2; Length 296;
Best Local Similarity 73.4%; Pred. No. 2.2e-24;
Matches 171; Conservative 0; Mismatches 56; Indels 6; Gaps 1;
```

QY 33 ATGGCTAGCATGACTGGTGACAGCAAAATGGGTGGGATCCTGA-----CCTGGAACAA 86
 Db 1 ATGGCTAGCATGACTGGTGACAGCAAAATGGGTGGGATCCTGGCGGTGGCCCTGGGAACAA 60
 QY 87 CGTGTGAGAACTGGCCCTGAAACGAGAGCTGGAAAAGAGAGCCCGAAGAGCTGGAG 146
 Db 61 CGTGTGAGAACTGGCCCTGAAACGAGAGCTGGAAAAGAGAGCCCGAAGAGCTGGAG 120
 QY 147 CAGGAAACCTGAACTCGAGGCGGAGACAGACCAACTAGAAATGAGAAGTCTGCTTTG 206
 Db 121 CAGGAAACCTGAACTCGAGGCGGAGAGTGTGGAGTTGGAAGTCTGTAATGACCCCTG 180
 QY 207 CAGACCGAGATTGCCAATCTGCTGAAGGAGAAAGAAAAGTGAAGTTTATCTT 259
 Db 181 CGCAAGGAGTGAACAGCTGGAGCGTGAACCTGGACACGCTGGCGGGTATCTT 233

RESULT 15

US-09-299-495F-32
 ; Sequence 32, Application US/09299495F
 ; Patent No. 6361968
 ; GENERAL INFORMATION:
 ; APPLICANT: VINSON, Charles R.
 ; KRYLOV, Dmitry
 ; TITLE OF INVENTION: EXTENSION OF A PROTEIN-PROTEIN
 ; INTERACTION SURFACE TO INACTIVATE THE FUNCTION OF A
 ; CELLULAR PROTEIN
 ; NUMBER OF SEQUENCES: 64
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: MORGAN & FINNEGAN, L.L.P.
 ; STREET: 345 Park Avenue
 ; CITY: New York
 ; STATE: NY
 ; COUNTRY: USA
 ; ZIP: 10154-0053
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: MS WORD 97
 ; SOFTWARE: ASCII
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/299,495F
 ; FILING DATE: 26-Apr-1999
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 60/001,654
 ; FILING DATE: 31-JUL-1995
 ; APPLICATION NUMBER: 60/018,496
 ; FILING DATE: 29-MAY-1996
 ; APPLICATION NUMBER: 08/690,011
 ; FILING DATE: 31-JULY-1996
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Serunian, Leslie A.
 ; REGISTRATION NUMBER: 35,353
 ; REFERENCE/DOCKET NUMBER: 2026-4199US2
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (212)758-4800
 ; TELEFAX: (212)751-6849
 ; INFORMATION FOR SEQ ID NO: 32:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 296 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: unknown
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: DNA (genomic)
 ; SEQUENCE DESCRIPTION: SEQ ID NO: 32:
 US-09-299-495F-32

Query Match 40.5%; Score 121.4; DB 4; Length 296;
 Best Local Similarity 73.4%; Pred. No. 2.2e-24;
 Matches 171; Conservative 0; Mismatches 56; Indels 6; Gaps 1;
 QY 33 ATGGCTAGCATGACTGGTGACAGCAAAATGGGTGGGATCCTGA-----CCTGGAACAA 86

Db 1 ATGGCTAGCATGACTGGTGACAGCAAAATGGGTGGGATCCTGGCGGTGGCCCTGGGAACAA 60
 QY 87 CGTGTGAGAACTGGCCCTGAAACGAGAGCTGGAAAAGAGAGCCCGAAGAGCTGGAG 146
 Db 61 CGTGTGAGAACTGGCCCTGAAACGAGAGCTGGAAAAGAGAGCCCGAAGAGCTGGAG 120
 QY 147 CAGGAAACCTGAACTCGAGGCGGAGACAGACCAACTAGAAATGAGAAGTCTGCTTTG 206
 Db 121 CAGGAAACCTGAACTCGAGGCGGAGAGTGTGGAGTTGGAAGTCTGTAATGACCCCTG 180
 QY 207 CAGACCGAGATTGCCAATCTGCTGAAGGAGAAAGAAAAGTGAAGTTTATCTT 259
 Db 181 CGCAAGGAGTGAACAGCTGGAGCGTGAACCTGGACACGCTGGCGGGTATCTT 233

Search completed: November 16, 2003, 00:58:53
 Job time : 51 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model
Run on: November 15, 2003, 23:56:10 : Search time 1512 Seconds
(without alignments)
8116.989 Million cell updates/sec

Title: US-10-059-720-18
Perfect score: 300
Sequence: 1 CCATGACTACAGGACGAC.....CAAGATCCCTGATTAGCTT 300

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2888711 seqs, 2045481386 residues
Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenBank

- 1: gb_ba.*
- 2: gb_hg.*
- 3: gb_in.*
- 4: gb_om.*
- 5: gb_ov.*
- 6: gb_pat.*
- 7: gb_ph.*
- 8: gb_pi.*
- 9: gb_pr.*
- 10: gb_ro.*
- 11: gb_sts.*
- 12: gb_sv.*
- 13: gb_un.*
- 14: gb_vi.*
- 15: em_ba.*
- 16: em_fun.*
- 17: em_hum.*
- 18: em_in.*
- 19: em_ma.*
- 20: em_on.*
- 21: em_or.*
- 22: em_ov.*
- 23: em_pat.*
- 24: em_ph.*
- 25: em_pl.*
- 26: em_ro.*
- 27: em_sts.*
- 28: em_un.*
- 29: em_vi.*
- 30: em_hcg_hum.*
- 31: em_hcg_inv.*
- 32: em_hcg_other.*
- 33: em_hcg_mus.*
- 34: em_hcg_pln.*
- 35: em_hcg_rtd.*
- 36: em_hcg_nam.*
- 37: em_hcg_vrt.*
- 38: em_sy.*
- 39: em_hgo_hum.*
- 40: em_hgo_mus.*
- 41: em_hgo_other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	300	100.0	300	6	AR202098	Sequence
2	267.2	89.1	281	6	AR202097	Sequence
3	179.4	59.8	324	6	AR202096	Sequence
4	160.6	53.5	266	6	AR202094	Sequence
5	133.4	44.5	262	6	AR202095	Sequence
6	130.8	43.6	1823	9	BC004490	Homo sapi
7	130.8	43.6	1918	9	AK097379	Homo sapi
8	130.8	43.6	2084	6	AX587848	Sequence
9	130.8	43.6	2103	6	AX014320	Sequence
10	130.2	43.4	273	6	AX098479	Sequence
11	124.4	41.5	700	9	AB022276	Homo sapi
12	124.4	41.5	2015	4	AF540379	Felis cat
13	124.4	41.5	3565	6	AX330284	Sequence
14	124.4	41.5	3565	6	AX336394	Sequence
15	124.4	41.5	3565	6	AX663638	Sequence
16	124.4	41.5	3565	6	I96207	Sequence
17	124.4	41.5	3565	9	H8CFOS	Human cellu
18	124.4	41.5	5897	9	AY212879	Homo sapi
19	124.4	41.5	6210	6	AX663628	Sequence
20	124.4	41.5	6210	9	HUMFOS	Human fos p
21	124.4	41.5	181848	9	CNS07YOR	Human chr
22	124.4	41.5	202367	9	DJ293M10	Human chr
23	124.4	41.5	207930	2	AC120203	Pan trogl
24	122.8	40.9	195254	2	AC117934	Papio anu
25	122.8	40.9	197796	2	AC117932	Papio anu
26	121.4	40.5	296	6	AR202105	Sequence
27	119.8	39.9	228	4	ORF535325	Ovis arie
28	119.8	39.9	140332	2	AC118538	Felis cat
29	119.6	39.9	135050	2	AC118536	Canis fam
30	119.6	39.3	381	4	AF069515	Bos tauru
31	116.4	38.8	414	4	SSCFOS34	Bos tauru
32	116.4	38.8	4200	4	SSC132510	Sus scrofa
33	116.4	38.8	188938	2	AC119064	Sus scrof
34	113.4	37.8	414	4	OAY15747	Ovis arie
35	113.2	37.7	174298	2	AC120514	Bos tauru
36	112.4	37.5	300	6	AX396221	Sequence
37	111.6	37.2	418	6	BD169720	C-terminu
38	111.6	37.2	2107	10	BC029814	Mus muscu
39	111.6	37.2	3811	14	AF033814	Murine os
40	111.6	37.2	3811	14	MSVMDSV	FBR murine
41	108.4	36.1	4226	14	REMSV5	Provirus of
42	106.8	35.6	1146	10	AF033012	Cricetulu
43	106.8	35.6	2116	10	RNCFOSR	Rat c-fos m
44	105.2	35.1	3673	10	MUSFOS	Mouse c-fos
45	105.2	35.1	3967	6	AX306224	Sequence

ALIGNMENTS

RESULT 1
AR202098
LOCUS AR202098
DEFINITION Sequence 18 from patent US 6361968.
ACCESSION AR202098
VERSION AR202098.1
KEYWORDS GI:20256637
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 300)
AUTHORS Winslow, C.R. and Krylov, D.
TITLE Extension of a protein-protein interaction surface to inactive the function of a cellular protein
JOURNAL Patent: US 6361968-A 18 26-MAR-2002;

AR202098 300 bp DNA linear PAT 20-APR-2002

```
FEATURES
  source
BASE COUNT    96 a    68 c    89 g    47 t
ORIGIN
Query Match      100.0%; Score 300; DB 6; Length 300;
Best Local Similarity 100.0%; Pred. No. 4.6e-61;
Matches 300; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CCATGGACTCAAGGACGAGATGACAAAGCATATGGCTAGCATGACTGTGGACAGCAA 60
DB 1 CCATGGACTCAAGGACGAGATGACAAAGCATATGGCTAGCATGACTGTGGACAGCAA 60
QY 61 TGGTCGGGATCTCTGACCTGGAAACACGCTGTGAGGAACCTGGCCCGTGAAGAGAGC 120
DB 61 TGGTCGGGATCTCTGACCTGGAAACACGCTGTGAGGAACCTGGCCCGTGAAGAGAGC 120
QY 121 TGGAAAAGAGGCGGAAAGCTGGAGCAGAGAAACGCTGAACCTCGAGGCGGAGACACC 180
DB 121 TGGAAAAGAGGCGGAAAGCTGGAGCAGAGAAACGCTGAACCTCGAGGCGGAGACACC 180
QY 181 AACTAGAGATGAGAACTGCTTTGCAGACCGAGATTGCCAACCTGCTCAAGGAGAGG 240
DB 181 AACTAGAGATGAGAACTGCTTTGCAGACCGAGATTGCCAACCTGCTCAAGGAGAGG 240
QY 241 AAAAAGTAGAGTTCATCTCGGAGCTCAGCAGCTCAGCAGCTGCTGCAAGATCCCTGATTAA 300
DB 241 AAAAAGTAGAGTTCATCTCGGAGCTCAGCAGCTCAGCAGCTGCTGCAAGATCCCTGATTAA 300

RESULT 2
LOCUS      AR202097      281 bp      DNA      linear      PAT 20-APR-2002
DEFINITION Sequence 16 from patent US 6361968.
ACCESSION  AR202097
VERSION     AR202097.1  GI:20256636
KEYWORDS    Unknown.
SOURCE      Unknown.
ORGANISM    Unclassified.
REFERENCE   1 (bases 1 to 281)
AUTHORS    Vinson,C.R. and Krylov,D.
TITLE      Extension of a protein-protein interaction surface to inactive the
JOURNAL    Patent: US 6361968-A 16 28-MAR-2002;
FEATURES    Location/Qualifiers
            source
            1..281
            /organism="unknown"
BASE COUNT    90 a    62 c    81 g    48 t
ORIGIN
Query Match      89.1%; Score 267.2; DB 6; Length 281;
Best Local Similarity 98.9%; Pred. No. 2.9e-53;
Matches 269; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 25 ACAAGCATATGGCTAGCATGACTGCTGGACAGCAATGGGTGGGATCTGACTGGAAC 84
DB 1 ATATACATATGGCTAGCATGACTGCTGGACAGCAATGGGTGGGATCTGACTGGAAC 60
QY 85 AACGTGCTGAGGAACCTGGCCCGTGAAGAACGAGAGCTGGAAAGAGGCGGAGAGCTGG 144
DB 61 AACGTGCTGAGGAACCTGGCCCGTGAAGAACGAGAGCTGGAAAGAGGCGGAGAGCTGG 120
QY 145 AGCAGGAACCGCTGAAGCTGAGCGGAGAGCAGACCACTAGAGATGAGAGTCTGCTT 204
DB 121 AGCAGGAACCGCTGAAGCTGAGCGGAGAGCAGACCACTAGAGATGAGAGTCTGCTT 180
QY 205 TGCAGACCGAGATTGCAACCTGCTGAAGGAGAGGAAACCTAGAGTTTCATCTCGCAG 264
DB 181 TGCAGACCGAGATTGCAACCTGCTGAAGGAGAGGAAACCTAGAGTTTCATCTCGCAG 240
QY 265 CTCACCGACCTGCTGCAAGATCCCTGATTAA 296

FEATURES
  source
BASE COUNT    96 a    68 c    89 g    47 t
ORIGIN
Query Match      100.0%; Score 300; DB 6; Length 300;
Best Local Similarity 100.0%; Pred. No. 4.6e-61;
Matches 300; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CCATGGACTCAAGGACGAGATGACAAAGCATATGGCTAGCATGACTGTGGACAGCAA 60
DB 1 CCATGGACTCAAGGACGAGATGACAAAGCATATGGCTAGCATGACTGTGGACAGCAA 60
QY 61 TGGTCGGGATCTCTGACCTGGAAACACGCTGTGAGGAACCTGGCCCGTGAAGAGAGC 120
DB 61 TGGTCGGGATCTCTGACCTGGAAACACGCTGTGAGGAACCTGGCCCGTGAAGAGAGC 120
QY 121 TGGAAAAGAGGCGGAAAGCTGGAGCAGAGAAACGCTGAACCTCGAGGCGGAGACACC 180
DB 121 TGGAAAAGAGGCGGAAAGCTGGAGCAGAGAAACGCTGAACCTCGAGGCGGAGACACC 180
QY 181 AACTAGAGATGAGAACTGCTTTGCAGACCGAGATTGCCAACCTGCTCAAGGAGAGG 240
DB 181 AACTAGAGATGAGAACTGCTTTGCAGACCGAGATTGCCAACCTGCTCAAGGAGAGG 240
QY 241 AAAAAGTAGAGTTCATCTCGGAGCTCAGCAGCTCAGCAGCTGCTGCAAGATCCCTGATTAA 300
DB 241 AAAAAGTAGAGTTCATCTCGGAGCTCAGCAGCTCAGCAGCTGCTGCAAGATCCCTGATTAA 300

RESULT 3
LOCUS      AR202096      324 bp      DNA      linear      PAT 20-APR-2002
DEFINITION Sequence 14 from patent US 6361968.
ACCESSION  AR202096
VERSION     AR202096.1  GI:20256635
KEYWORDS    Unknown.
SOURCE      Unknown.
ORGANISM    Unclassified.
REFERENCE   1 (bases 1 to 324)
AUTHORS    Vinson,C.R. and Krylov,D.
TITLE      Extension of a protein-protein interaction surface to inactive the
JOURNAL    Patent: US 6361968-A 14 26-MAR-2002;
FEATURES    Location/Qualifiers
            source
            1..324
            /organism="unknown"
BASE COUNT    109 a    71 c    93 g    51 t
ORIGIN
Query Match      59.8%; Score 179.4; DB 6; Length 324;
Best Local Similarity 77.8%; Pred. No. 2.2e-32;
Matches 253; Conservative 0; Mismatches 46; Indels 26; Gaps 2;
QY 1 CCATGGACTCAAGGACGAGATGACAAAGCATATGGCTAGCATGACTGTGGACAGCAA 60
DB 1 CCATGGACTCAAGGACGAGATGACAAAGCATATGGCTAGCATGACTGTGGACAGCAA 60
QY 61 TGGTCGGGATCTCTGACCTGGAAACACGCTGTGAGGAACCTGGCCCGTGAAGAGAGC 120
DB 61 TGGTCGGGATCTCTGACCTGGAAACACGCTGTGAGGAACCTGGCCCGTGAAGAGAGC 119
QY 121 TGGAAAAGAGGCGGAAAGCTGGAGCAGAGAAACCTAGAGTTTCATCTCGCAGCTCACC 155
DB 120 CGAAGGGAAGGAATAGATGGCTGCAGCCAAATGCCGCAACCGGAGGAGGAGTGA 179
QY 156 GCTGTAAGCTGAGCGGAGACAGACCACTAGAGATGAGAGTCTCTTTGAGAGCCGAG 215
DB 180 GATACACTCCAGCGGAGACAGACCACTAGAGATGAGAGTCTCTTTGAGAGCCGAG 239
QY 216 ATTGCCAACCTGCTGAAGGAGAGGAAACCTAGAGTTTCATCTCGCAGCTCACCAGCT 275
DB 240 ATTGCCAACCTGCTGAAGGAGAGGAAACCTAGAGTTTCATCTCGCAGCTCACCAGCT 299
QY 276 GCTGCAAGATCTCTGATTAAAGCTT 300
DB 300 GCTGCAAGATCTCTGATTAAAGCTT 324

RESULT 4
LOCUS      AR202094      286 bp      DNA      linear      PAT 20-APR-2002
DEFINITION Sequence 10 from patent US 6361968.
ACCESSION  AR202094
VERSION     AR202094.1  GI:20256633
KEYWORDS    Unknown.
SOURCE      Unknown.
ORGANISM    Unclassified.
REFERENCE   1 (bases 1 to 286)
AUTHORS    Vinson,C.R. and Krylov,D.
TITLE      Extension of a protein-protein interaction surface to inactive the
JOURNAL    Patent: US 6361968-A 10 26-MAR-2002;
FEATURES    Location/Qualifiers
            source
            1..286
            /organism="unknown"
BASE COUNT    93 a    54 c    75 g    44 t
```

ORIGIN

Query Match 53.5%; Score 160.6; DB 6; Length 266;
Best Local Similarity 97.6%; Pred. No. 6.5e-28;
Matches 163; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 CCATGGACTCAAGGACGATGACAGCATATGGCTAGCATGCTGGTGGACAGCAA 60
DB 1 CCATGGACTCAAGGACGATGACAGCATATGGCTAGCATGCTGGTGGACAGCAA 60

QY 61 TGGTGGGATCTGACTGGAAACAAGTCTGAGAACTGGCCCGTGAAGAACAGAGGC 120
DB 61 TGGTGGGATCTGACTGGAAACAAGTCTGAGAACTGGCCCGTGAAGAACAGAGGC 120

QY 121 TGGAAAGAGGCGGAGAGCTGAGCAGGAGAAACGCTGAACCTGAG 167
DB 121 TGGAAAGAGGCGGAGAGCTGAGCAGGAGAAACGCTGAGAACTGAG 167

RESULT 5
AR202095
LOCUS AR202095 Sequence 12 from patent US 6361968.
DEFINITION AR202095
ACCESSION AR202095.1 GI:20256634
VERSION AR202095.1
KEYWORDS Unknown.
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 262)
AUTHORS Vinson,C.R. and Krylov,D.
TITLE Extension of a protein-protein interaction surface to inactive the function of a cellular protein
JOURNAL Patent: US 6361968-A 12 26-MAR-2002;
FEATURES Location/Qualifiers
SOURCE 1..262
BASE COUNT 90 a 58 c 72 g 42 t
ORIGIN

Query Match 44.5%; Score 133.4; DB 6; Length 262;
Best Local Similarity 92.7%; Pred. No. 1.9e-21;
Matches 140; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 146 GCAGGAAACGCTGAACTGAGGCGGAGACAGACCACTAGAGATGAGAGCTGCTTT 205
DB 103 GGAGCTGATGATACACTCCAGCGGAGACAGACCACTAGAGATGAGAGCTGCTTT 162

QY 206 GCAGACCGAGATTGCCAACCTGCTGAGGAGAGAGAGAGAGAGAGAGAGAGAG 265
DB 163 GCAGACCGAGATTGCCAACCTGCTGAGGAGAGAGAGAGAGAGAGAGAGAGAG 222

QY 266 TCACCGACTGCTGCAAGATCCCTGATTA 296
DB 223 TCACCGACTGCTGCAAGATCCCTGATTA 253

RESULT 6
BC004490
LOCUS BC004490 1823 bp mRNA linear PRI 12-JUL-2001
DEFINITION Homo sapiens, v-fos FBJ murine osteosarcoma viral oncogene homolog, clone MGC:11074 IMAGE:368870, mRNA, complete cds.
ACCESSION BC004490
VERSION BC004490.1 GI:13325363
KEYWORDS MGC.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS 1 (bases 1 to 1823)
TITLES Strausberg, R.
JOURNAL Direct Submission
Submitted (12-MAR-2001) National Institutes of Health, Mammalian

Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
NIH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Institute for Systems Biology
<http://www.systemsbio.org>
contact: amandare@systemsbiology.org
Anup Madan, Rachel Dickhoff, Jessica Fahey, Stephanie Ford, Julia Greene, Mark Kettelman and Anuradha Madan

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAL Plate: 14 Row: 0 Column: 17
This clone was selected for full length sequencing because it passed the following selection criteria: Hexamer frequency ORF analysis, similarity but not identity to protein.

FEATURES
Location/Qualifiers
1..1823
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="LocusID:2353"
/db_xref="taxon:9606"
/clone="MGC:11074 IMAGE:368870"
/tissue type="Pancreas, adenocarcinoma"
/clone_lib="NIH MGC 39"
/lab_host="DH10B-R"
/note="Vector: pOTF7"
136..1278
/codon_start=1
/product="v-fos FBJ murine osteosarcoma viral oncogene homolog"
/protein_id="AAH04490.1"
/db_xref="GI:13325364"
/translation="MFMGFNADYEAASSRCSSASPGDLSLYHSPADSESSMGSPV
NAQDFCTDLAVSSANFIPTVTAISTPDLQVLQVLPALVSSVAPQTRAPHFGVPAPS
AGAGAGVWKTMTGTRAGSIRGRKVEQLSPFEEERIRIRRNKAAACRRRRE
LTDITQATDOLEDEKSAIQETIANILAKKEKLEFLLAAHPACKIPDGLGFPBMSV
ASLDTLGLPEVATPESEARTPLINDPEPKPSVEPVKSISSMELKTEPDDFLFPA
SRPSSSTARSVPDMDSGSPYADWEPHLSGLGMPMATELEPLCTPVVTCPTSC
TATSSFTFTFEDRSPFCASAHKNGSSNEPSSDLSLSPFLLAL"

BASE COUNT 416 a 545 c 476 g 386 t
ORIGIN

Query Match 43.6%; Score 130.8; DB 9; Length 1823;
Best Local Similarity 92.0%; Pred. No. 9.1e-21;
Matches 138; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

QY 146 GCAGGAAACGCTGAACTGAGGCGGAGACAGACCACTAGAGATGAGAGCTGCTTT 205
DB 612 GCGAGCTGACTGATACACTCCAGCGGAGACAGACCACTAGAGATGAGAGCTGCTTT 671

QY 206 GCAGACCGAGATTGCCAACCTGCTGAGGAGAGAGAGAGAGAGAGAGAGAGAG 265
DB 672 GCAGACCGAGATTGCCAACCTGCTGAGGAGAGAGAGAGAGAGAGAGAGAGAG 731

QY 266 TCACCGACTGCTGCAAGATCCCTGATTA 295
DB 732 TCACCGACTGCTGCAAGATCCCTGATTA 761

RESULT 7
AK097379
LOCUS AK097379 1918 bp mRNA linear PRI 15-JUL-2002
DEFINITION Homo sapiens cDNA FLJ40060 fis, clone TCOLN200236, highly similar to P55-C-FOS PROTO-ONCOGENE PROTEIN.
ACCESSION AK097379
VERSION AK097379.1 GI:21757114
KEYWORDS oligo capping; fis (full insert sequence).

```

SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
REFERENCE   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS     Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
Kodaikiya, K., Wagatsuma, M., Kanda, K., Kondo, H., Yokoi, T.,
Kamihara, K., Katsuta, N., Sato, K., Tanikawa, M., Yamazaki, M.,
Sugiyama, T., Irie, R., Otsuki, T., Sato, H., Wakamatsu, A., Ishii, S.,
Yamamoto, J., Isono, Y., Kawai-Hio, Y., Saito, K., Nishikawa, T.,
Kimura, K., Yamashita, H., Matsuo, K., Nakamura, Y., Sekine, M.,
Kikuchi, H., Murakawa, K., Kanehori, K., Takahashi-Fujii, A.,
Oshima, A., Sugiyama, A., Kawakami, S., Suzuki, Y., Sugano, S.,
Nagahara, K., Masuko, Y., Nagai, K. and Isogai, T.
TITLE      NEDO human cDNA sequencing project
JOURNAL     Unpublished
REFERENCE   2 (bases 1 to 1918)
AUTHORS     Isogai, T. and Yamamoto, J.
TITLE      Direct Submission
JOURNAL     Submitted (04-JUL-2002) Takao Isogai, FLJ Project (HRI Team); 2-6-7
Karusa-Kamatari, Kisarazu, Chiba 292-0812, Japan
(Email: genomics@hri.co.jp, Tel: 81-438-52-3975, Fax: 81-438-52-3986)
COMMENT     NEDO human cDNA sequencing project supported by Ministry of
Economy, Trade and Industry of Japan; cDNA full insert sequencing:
Research Association for Biotechnology (RAB); cDNA library
construction: Helix Research Institute (HRI) (supported by Japan
Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB,
HRI, and Biotechnology Center, National Institute of Technology and
Evaluation; clone selection for full insert sequencing: HRI and
RAB; annotation: HRI and RAB.
FEATURES   Location/Qualifiers
            1..1918
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /clone="TCOLN2000236"
            /tissue_type="colon, tumor tissue"
            /clone_lib="TCOLN2"
            /note="Cloning vector: pME18SFL3"
BASE COUNT 452 a 505 c 482 g 479 t
ORIGIN
Query Match 43.6%; Score 130.8; DB 9; Length 1918;
Best Local Similarity 92.0%; Pred. No. 9.1e-21;
Matches 138; Conservative 0; Mismatches 12; Indels 0; Gaps 0;
QY 146 GCAGGAAACGCTGAACTCGAGGCGGAGACAGACCACTAGAGATGAGAAGTCTGCTTT 205
Db 447 GGAGCTGACTGATACACTCCAAAGCGGAGACAGACCACTAGAGATGAGAAGTCTGCTTT 506
QY 206 GCAGACCGAGATTGCCACCTGCTGAGGAGAGGAAAACCTAGAGTTTCATCTCGGCAGC 265
Db 507 GCAGACCGAGATTGCCACCTGCTGAGGAGAGGAAAACCTAGAGTTTCATCTCGGCAGC 566
QY 266 TCACCGACCTGCTGCAAGATCCCTGATTA 295
Db 567 TCACCGACCTGCTGCAAGATCCCTGATGA 596

RESULT 8
AX587848
LOCUS      2084 bp DNA linear PAT 10-JAN-2003
DEFINITION Sequence 318 from Patent WO0246467.
ACCESSION AX587848
VERSION   AX587848.1 GI:28212432
KEYWORDS  synthetic construct
SOURCE    synthetic construct
ORGANISM  artificial sequences.
REFERENCE 1
AUTHORS   Bertucci, F., Houlgatte, R., Birnbaum, D., Nguyen, C., Viens, P. and
Pert, V.
TITLE     Gene expression profiling of primary breast carcinomas using arrays

```

```
RESULT 10
AX098479/c
LOCUS
DEFINITION Sequence 16 from Patent WO0120029.
ACCESSION AX098479
VERSION AX098479.1 GI:13537765
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 Tocque, B., Bracco, L. and Schweighoffer, F.
AUTHORS Genetic markers of toxicity, preparation and uses thereof
TITLE Patent: WO 0120029-A 16 22-MAR-2001;
JOURNAL Exonhit Therapeutics S.A. (FR)
FEATURES
source
1..273
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/notes="Identity to human c-fos proto-oncogene exon 4 cds
nucleotides 2687-2908. GenBank Acc: K00650."
BASE COUNT 52 a 74 c 79 g 58 t
ORIGIN
Query Match 43.4%; Score 130.2; DB 6; Length 273;
Best Local Similarity 97.8%; Pred. No. 1.1e-20;
Matches 132; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 161 ACTCGGCGGAGACACCACTAGAGATGAGAGTCTGCTTGCAGACCGAGATTGC 220
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
237 ACTCCAGCGGAGACACCACTAGAGATGAGAGTCTGCTTGCAGACCGAGATTGC 178
QY 221 CAACCTGCTGAGAGAGAGAGAACTAGAGTTCATCTCGCAGCTCACCAGCTGCCTG 280
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
177 CAACCTGCTGAGAGAGAGAGAACTAGAGTTCATCTCGCAGCTCACCAGCTGCCTG 118
QY 281 CAAGATCCCTGATTA 295
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
117 CAAGATCCCTGATGA 103

RESULT 11
AB022276
LOCUS
DEFINITION Homo sapiens gene for cellular oncogene c-fos, partial cds.
ACCESSION AB022276
VERSION AB022276.1 GI:6518634
KEYWORDS cellular oncogene c-fos.
SOURCE Homo sapiens (human)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 700)
AUTHORS Umino, Y., Hohjoh, H. and Tokunaga, K.
TITLE Human c-fos gene variant(2) at the position of 3044 in the genomic
sequence (#V01512)
JOURNAL Published Only in DataBase (1999)
REFERENCE
2 (bases 1 to 700)
AUTHORS Umino, Y., Hohjoh, H. and Tokunaga, K.
TITLE Direct Submission
JOURNAL Submitted (11-JAN-1999) Yumiko Umino, University of Tokyo,
Department of Human Genetics; 7-3-1 Hongo, Bunkyo-ku, Tokyo Japan,
Bunkyo-ku, Tokyo 113-0033, Japan (E-mail: Y-umino@u-tokyo.ac.jp,
Tel: 81-3-3812-2111 (ex.3654), Fax: 81-3-5802-8619)
FEATURES
source
1..700
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/notes="Identity to human c-fos proto-oncogene exon 4 cds
nucleotides 2687-2908. GenBank Acc: K00650."
BASE COUNT 52 a 74 c 79 g 58 t
ORIGIN
Query Match 43.4%; Score 130.2; DB 6; Length 273;
Best Local Similarity 97.8%; Pred. No. 1.1e-20;
Matches 132; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 161 ACTCGGCGGAGACACCACTAGAGATGAGAGTCTGCTTGCAGACCGAGATTGC 220
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
237 ACTCCAGCGGAGACACCACTAGAGATGAGAGTCTGCTTGCAGACCGAGATTGC 178
QY 221 CAACCTGCTGAGAGAGAGAGAACTAGAGTTCATCTCGCAGCTCACCAGCTGCCTG 280
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
177 CAACCTGCTGAGAGAGAGAGAACTAGAGTTCATCTCGCAGCTCACCAGCTGCCTG 118
QY 281 CAAGATCCCTGATTA 295
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
117 CAAGATCCCTGATGA 103

RESULT 12
AF540379
LOCUS
DEFINITION Felis catus immediate early protein (c-fos) mRNA, complete cds.
ACCESSION AF540379
VERSION AF540379.2 GI:23346378
KEYWORDS
SOURCE Felis catus (cat)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Carnivora; Fissipedia; Felidae; Felis.
REFERENCE
1 (bases 1 to 2015)
AUTHORS Van der Gucht, E., Massie, A., De Klerck, B., Peeters, K., Winters, K.,
Gerets, H., Clerens, S., Vandesande, F. and Arckens, L.
TITLE Molecular cloning and differential expression of the cat immediate
early gene c-fos
JOURNAL Brain Res. Mol. Brain Res. 111 (1-2), 198-210 (2003)
REFERENCE
2 (bases 1 to 2015)
AUTHORS Van der Gucht, E., Massie, A., De Klerck, B., Peeters, K., Winters, K.,
Gerets, H., Clerens, S., Vandesande, F. and Arckens, L.
TITLE Direct Submission
JOURNAL Submitted (22-AUG-2002) Biology, K. U. Leuven, Naamsestraat 59,
Leuven B-3000, Belgium
REFERENCE
3 (bases 1 to 2015)
AUTHORS Van der Gucht, E., Massie, A., De Klerck, B., Peeters, K., Winters, K.,
Gerets, H., Clerens, S., Vandesande, F. and Arckens, L.
TITLE Direct Submission
JOURNAL Submitted (30-SEP-2002) Biology, K. U. Leuven, Naamsestraat 59,
Leuven B-3000, Belgium
REMARK Sequence update by submitter
COMMENT On Sep 30, 2002 this sequence version replaced gi:23267178.
FEATURES
source
1..2015
/db_xref="taxon:9606"
/chromosome="14"
<1..37
/number=3
38..>700
/number=4
<38..679
/codon_start=1
/product="cellular oncogene c-fos"
/protein_id="BAA87921.1"
/db_xref="GI:6518635"
/translation="ETDQLEDEKSLQTEIANLLKEKELEFLAAHRPAKIPDDLG
PPEKSVASLDLTGGLPEVATESSEAFPLPLNDPEPKSPVPEPKYSISIMELKTEPFP
DQFLFPASSRPSGSETARSPVDMDSGFYAADWELPHSGSLMGPMATELEPLCTFPV
VTCPTSCATYTSFVFTYPEADSPFSCAAHRRKSSSNFSDSUSPTLLAL"
variation
334
/replaces="c"
BASE COUNT 140 a 225 c 188 g 147 t
ORIGIN
Query Match 41.5%; Score 124.4; DB 9; Length 700;
Best Local Similarity 99.2%; Pred. No. 2.8e-19;
Matches 125; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 170 GGAGACAGACCACTAGAGATGAGAGTCTGCTTGCAGACCGAGATTGCCAACCTGCT 229
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
37 GGAGACAGACCACTAGAGATGAGAGTCTGCTTGCAGACCGAGATTGCCAACCTGCT 96
QY 230 GGAGAGAGAGAGAGAGAGTTCATCTCGCAGCTCACCAGCTGCCTGCAAGATCCC 289
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
97 GGAGAGAGAGAGAGAGAGTTCATCTCGCAGCTCACCAGCTGCCTGCAAGATCCC 156
QY 290 TGATTA 295
Db | | | | |
157 TGATGA 162

RESULT 13
AF540379
LOCUS
DEFINITION Felis catus immediate early protein (c-fos) mRNA, complete cds.
ACCESSION AF540379
VERSION AF540379.2 GI:23346378
KEYWORDS
SOURCE Felis catus (cat)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Carnivora; Fissipedia; Felidae; Felis.
REFERENCE
1 (bases 1 to 2015)
AUTHORS Van der Gucht, E., Massie, A., De Klerck, B., Peeters, K., Winters, K.,
Gerets, H., Clerens, S., Vandesande, F. and Arckens, L.
TITLE Molecular cloning and differential expression of the cat immediate
early gene c-fos
JOURNAL Brain Res. Mol. Brain Res. 111 (1-2), 198-210 (2003)
REFERENCE
2 (bases 1 to 2015)
AUTHORS Van der Gucht, E., Massie, A., De Klerck, B., Peeters, K., Winters, K.,
Gerets, H., Clerens, S., Vandesande, F. and Arckens, L.
TITLE Direct Submission
JOURNAL Submitted (22-AUG-2002) Biology, K. U. Leuven, Naamsestraat 59,
Leuven B-3000, Belgium
REFERENCE
3 (bases 1 to 2015)
AUTHORS Van der Gucht, E., Massie, A., De Klerck, B., Peeters, K., Winters, K.,
Gerets, H., Clerens, S., Vandesande, F. and Arckens, L.
TITLE Direct Submission
JOURNAL Submitted (30-SEP-2002) Biology, K. U. Leuven, Naamsestraat 59,
Leuven B-3000, Belgium
REMARK Sequence update by submitter
COMMENT On Sep 30, 2002 this sequence version replaced gi:23267178.
FEATURES
source
1..2015
/db_xref="taxon:9606"
/chromosome="14"
<1..37
/number=3
38..>700
/number=4
<38..679
/codon_start=1
/product="cellular oncogene c-fos"
/protein_id="BAA87921.1"
/db_xref="GI:6518635"
/translation="ETDQLEDEKSLQTEIANLLKEKELEFLAAHRPAKIPDDLG
PPEKSVASLDLTGGLPEVATESSEAFPLPLNDPEPKSPVPEPKYSISIMELKTEPFP
DQFLFPASSRPSGSETARSPVDMDSGFYAADWELPHSGSLMGPMATELEPLCTFPV
VTCPTSCATYTSFVFTYPEADSPFSCAAHRRKSSSNFSDSUSPTLLAL"
variation
334
/replaces="c"
BASE COUNT 140 a 225 c 188 g 147 t
ORIGIN
Query Match 41.5%; Score 124.4; DB 9; Length 700;
Best Local Similarity 99.2%; Pred. No. 2.8e-19;
Matches 125; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 170 GGAGACAGACCACTAGAGATGAGAGTCTGCTTGCAGACCGAGATTGCCAACCTGCT 229
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
37 GGAGACAGACCACTAGAGATGAGAGTCTGCTTGCAGACCGAGATTGCCAACCTGCT 96
QY 230 GGAGAGAGAGAGAGAGTTCATCTCGCAGCTCACCAGCTGCCTGCAAGATCCC 289
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
97 GGAGAGAGAGAGAGAGTTCATCTCGCAGCTCACCAGCTGCCTGCAAGATCCC 156
QY 290 TGATTA 295
Db | | | | |
157 TGATGA 162
```



```
FEATURES                               Location/Qualifiers
source                                i. .3565
                                     /organism="Homo sapiens"
                                     /mol_type="genomic DNA"
                                     /db_xref="taxon:9606"
BASE COUNT      780 a   978 c   853 t
ORIGIN
Query Match      41.5%; Score 124.4; DB 6; Length 3565;
Best Local Similarity 99.2%; Pred. No. 3.2e-19;
Matches 125; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY      170 GGGGACAGACCAACTAGAGATGAGAGTCTGCTTTGCAGACCGAGATTGCCAACCTGCT 229
Db      2087 GGAGACAGACCAACTAGAGATGAGAGTCTGCTTTGCAGACCGAGATTGCCAACCTGCT 2146
QY      230 GAAGGAGAGAGGAAAACTAGAGTTCATCTGGCAGCTCACCGACCTGCCTGCAAGATCCC 289
Db      2147 GAAGGAGAGGAAAACTAGAGTTCATCTGGCAGCTCACCGACCTGCCTGCAAGATCCC 2206
QY      290 TGATTA 295
Db      2207 TGATGA 2212
```

Search completed: November 16, 2003, 00:35:55
Job time : 1516 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: November 16, 2003, 01:30:16 ; Search time 1717 Seconds
(without alignments)
1373.054 Million cell updates/sec

Title: US-10-059-720-19

Perfect score: 493

Sequence: 1 MDYKDDDDKHWSMTGGQM.....EKEKLEFLAHRPACKIPD 97

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US10059720/runat_14112003_184959_17898/app_query.fasta_1.263
-DB=EST -QFMT=fastap -SUFFIX=1st -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=biosum62 -TRANS=human40.cdi -LIST=45
-DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=plo -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10059720@cgn.1.1.2810 @runat_14112003_184959_17898 -NCPU=6 -ICPU=3
-NO WMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

EST.*
1: em_estba.*
2: em_esthum.*
3: em_estin.*
4: em_estmu.*
5: em_estov.*
6: em_estpl.*
7: em_estro.*
8: em_hrc.*
9: gb_est1.*
10: gb_est2.*
11: gb_est3.*
12: gb_est4.*
13: gb_est5.*
14: gb_estfun.*
15: em_estom.*
16: em_estom.*
17: em_gss_hum.*
18: em_gss_inv.*
19: em_gss_pln.*
20: em_gss_vrt.*
21: em_gss_fun.*
22: em_gss_mam.*
23: em_gss_mus.*
24: em_gss_pro.*
25: em_gss_rod.*
26: em_gss_phg.*
27: em_gss_vrl.*
28: gb_gss1.*

29: gb_gss2.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description	
C	1	243	49.3	435	12	BG979535	BG979535 CM4-CN006
	2	241	48.9	428	9	AU297079	AU297079 AU297079
	3	241	48.9	550	9	AL703189	AL703189 DKF2P686C
	4	241	48.9	551	13	BQ694927	BQ694927 1001070 H
	5	241	48.9	565	14	CA406906	CA406906 1003055 H
	6	241	48.9	614	13	BQ694699	BQ694699 1000842 H
	7	241	48.9	615	12	BG924297	BG924297 HNC26-1-C
	8	241	48.9	635	13	BQ694763	BQ694763 1000906 H
	9	241	48.9	637	12	BM818004	BM818004 K-EST0084
	10	241	48.9	650	13	BQ693998	BQ693998 1000141 H
	11	241	48.9	654	12	BM820048	BM820048 K-EST0088
	12	241	48.9	667	12	BI481329	BI481329 H2RPE-161
	13	241	48.9	691	12	BM685886	BM685886 UI-B-CQ0
	14	241	48.9	713	12	BG925313	BG925313 HNC40-1-A
	15	241	48.9	755	14	CA777910	CA777910 ip15b02.x
	16	241	48.9	778	14	CA772039	CA772039 iO94G05.x
	17	241	48.9	849	14	CD107570	CD107570 AGENCOURT
C	18	241	48.9	895	13	BQ717306	BQ717306 AGENCOURT
	19	241	48.9	923	10	BG742465	BG742465 602635387
	20	241	48.9	975	13	BQ882100	BQ882100 AGENCOURT
	21	241	48.9	1139	12	BM544094	BM544094 AGENCOURT
C	22	240	48.7	288	14	CA388732	CA388732 C802C06.Y
	23	238	48.3	537	12	BM089359	BM089359 503083 MA
	24	236	47.9	561	10	BF853756	BF853756 MR2-BN009
	25	235	47.7	445	9	AK462661	AK462661 BP230010A
	26	235	47.7	724	13	BU487401	BU487401 604127017
	27	234	47.5	582	10	BF901154	BF901154 PM4-MT023
	28	233	47.3	603	14	CB554541	CB554541 MMSP0017
	29	233	47.3	632	14	BY729407	BY729407 BY729407
	30	232	47.1	539	14	CB613051	CB613051 ANGNNUC:U
	31	232	47.1	557	10	BE840523	BE840523 RCI-FN018
C	32	230	46.7	396	12	BQ192970	BQ192970 PM3-MT020
	33	229	46.5	326	13	BQ082970	BQ082970 K-EST0144
	34	229	46.5	439	10	BF724668	BF724668 bx07e01.Y
	35	229	46.5	703	10	BE373426	BE373426 601225088
	36	229	46.5	1129	13	BQ723263	BQ723263 AGENCOURT
C	37	229	46.5	1140	13	EX398353	EX398353 BX398353
	38	228	46.2	794	12	BI760850	BI760850 603043701
	39	228	46.2	912	13	BQ716024	BQ716024 AGENCOURT
C	40	228	46.2	976	12	BM545725	BM545725 AGENCOURT
	41	226	45.8	293	13	BU788894	BU788894 i192b11.Y
	42	226	45.8	421	14	CB764157	CB764157 ANGNNUC:N
	43	226	45.8	553	10	BE840522	BE840522 RCI-FN018
	44	226	45.8	601	13	BU696599	BU696599 L02in1221
45	226	45.8	789	10	BG695872	BG695872 602657836	

ALIGNMENTS

RESULT 1
LOCUS BG979535/C
DEFINITION CM4-CN0062-120101-794-b08 CN0062 Homo sapiens cDNA, mRNA linear EST 12-JUN-2001
ACCESSION BG979535
VERSION BG979535.1 GI:14382270
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Euthera; Primates; Cetarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 435)
AUTHORS Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,

BG979535 435 bp mRNA linear EST 12-JUN-2001
CM4-CN0062-120101-794-b08 CN0062 Homo sapiens cDNA, mRNA sequence.
BG979535
BG979535.1 GI:14382270

TITLE
Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F., Goldman, G.H., Carvalho, A.F., Matsukuma, A., Baia, G.S., Simpson, D.H., Brunstein, A., de Oliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J., and Simpson, A.J.

JOURNAL
Shotgun sequencing of the human transcriptome with ORF expressed sequence tags

MEDLINE
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

PUBMED
10737800

COMMENT
Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?ti=CM4&t2=CM4-CN0062-120101-794-b08&t3=2001-01-13&t4=1)
Seq primer: puc 18 forward
High quality sequence start: 44
High quality sequence stop: 431.

FEATURES source

1. 435
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="Adult"
/clone_lib="CN0062"
/notes="Organ: colon normal; Vector: puc18; Site 1: SmaI; Site 2: SmaI; A mini-library was made by cloning products derived from ORFESTES PCR (U.S. Letters Patent application No.156,716 - Ludwig Institute for Cancer Research) profiles into the puc 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

BASE COUNT 70 a 127 c 120 g 117 t 1 others

Alignment Scores:

Pred. No.: 1.88e-15 **Length:** 435
Score: 243.00 **Matches:** 55
Percent Similarity: 66.67% **Conservative:** 9
Best Local Similarity: 57.29% **Mismatches:** 20
Query Match: 49.29% **Indels:** 12
DB: 12 **Gaps:** 3

US-10-059-720-19 (1-97) x BG979535 (1-435)

QY 13 SerMetThrGlyGly-----GlnGlnMetGlyArgAspProAspLeuGluGlnArgAla 30
Db 310 ACCATGACAGGAGGCGCGAGCTCAGACGATTCGAGAGGCGGCGAGGTGAAACAGTATCT 251
QY 31 GluGluLeuAlaArgGluGlnGluGluGluGluGluGluGluGluGluGluGluGluGlu 50
Db 250 CCRAGAA---GAAGAGAGAGAAAGAGAGATCCGAAGGGAAGGAAATAGATGCTGCGAGCC 194
QY 51 AsnAlaGlu-----LeuGluAlaGluThrAspGlnLeu 61
Db 193 AATGCGCGNACCGAGGAGGAGCTGATACACTCCAGCGAGAGAGAGAGAGAGAGAGAG 134
QY 62 GluAspGluLysSerAlaLeuGlnThrGluLeuAlaAsnLeuLysGluLysGluLys 81
Db 133 GAAGATGAGAGAGCTCTCTTTGACAGCGAGATTCGCAACTGCTGAGAGAGAGAGAGAG 74
QY 82 LeuGluPheLeuAlaAlaHisArgProAlaCysLysIleProAsp 97
Db 73 CTAGAGTTATCTCTGCGAGCTCACCAGCTGCTGCAAGATACCTGAT 26

RESULT 2

AU297079
LOCUS AU297079 female adult cerebellum, full-length enriched chimpanzee
DEFINITION CDNA library Pan troglodytes verus cDNA clone FCB3064 5' similar to human RefSeq mRNA NM_005252, mRNA sequence.
ACCESSION AU297079
VERSION AU297079.1 GI:29531378
KEYWORDS EST.
SOURCE Pan troglodytes verus
ORGANISM Pan troglodytes verus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
REFERENCE 1 (bases 1 to 428)
AUTHORS Sakate, R., Osada, N., Hida, M., Sugano, S., Hayasaka, I., Shimohira, N., Yanagi, S., Suto, Y., Hashimoto, K. and Hirai, M.
ANALYSIS OF 5'-end sequences of chimpanzee cDNAs
Genome Res. 13 (5), 1022-1026 (2003)
JOURNAL 12727913
MEDLINE 12727913
PUBMED 12727913
COMMENT Contact: Momoki Hirai
Department of Integrated Biosciences
The University of Tokyo, Graduate School of Frontier Sciences
5-1-5 Kashiwanoha, Kashiwa, Chiba 277-8562, Japan
Tel: 81-4-7136-3688
Fax: 81-4-7136-3687
Email: mhira@u-tokyo.ac.jp.
Location/Qualifiers

FEATURES source

1. 428
/organism="Pan troglodytes verus"
/mol_type="mRNA"
/sub_species="verus"
/db_xref="taxon:37013"
/clone="PCCB3064"
/sex="female"
/tissue_type="cerebellum"
/dev_stage="adult"
/clone_lib="female adult cerebellum, full-length enriched chimpanzee cDNA library"

BASE COUNT 115 a 120 c 126 g 66 t 1 others

Alignment Scores:

Pred. No.: 2.95e-15 **Length:** 428
Score: 241.00 **Matches:** 57
Percent Similarity: 65.66% **Conservative:** 8
Best Local Similarity: 57.59% **Mismatches:** 16
Query Match: 48.88% **Indels:** 18
DB: 9 **Gaps:** 3

US-10-059-720-19 (1-97) x AU297079 (1-428)

QY 13 SerMetThrGlyGly-----GlnGlnMetGlyArgAspProAspLeuGluGlnArgAla 30
Db 124 ACCATGACAGGAGGCGCGAGCTCAGACGATTCGAGAGGCGGCGAGGTGAAACAGTATCT 171
QY 31 GluGluLeuAlaArgGluGlnGluGluGluGluGluGluGluGluGluGluGluGluGlu 50
Db 172 GAACAGTTATCTCCAG 231
QY 51 AsnAla-----GluLeuGluAlaGluThr 58
Db 232 GCTGCGAGCAAAATGCCGCAACCGAGGAGNAGCTGATACACTCCAGCGAGAGACA 291
QY 59 AspGlnLeuGluAspGluLysSerAlaLeuGlnThrGluLeuAlaAsnLeuLysGlu 78
Db 292 GACCACTAGAGATGAGAGATGCTGCTTTGACAGCCGAGATTCGCAACCTGCTGAGGAG 351
QY 79 LysGluLysLeuGluPheLeuAlaAlaHisArgProAlaCysLysIleProAsp 97
Db 352 AAGGAAAACTAGAGTTTCATCTCTGCGAGCTCACCAGCTGCTGCTGCAAGATCCCTGAT 408

RESULT 3
AL703189

LOCUS AL703189 550 bp mRNA linear EST 22-MAR-2002
 DEFINITION DKFZp686C2019.r1.686 (synonym: hlcc3) Homo sapiens cDNA clone
 accession DKFZp686C2019.5, mRNA sequence.
 VERSION AL703189.1 GI:19686544
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 550)
 AUTHORS Poustka, A., Wellenreuther, R., Mewes, H.W., Weil, B. and Wiemann, S.).
 TITLE EST (Poustka, A., Wellenreuther, R., Mewes, H.W., Weil, B. and Wiemann, S.).
 JOURNAL Unpublished
 COMMENT Contact: Poustka A.J.
 Department Lehrach
 Max-Planck-Institute for Molecular Genetics
 Ihnestrasse 73, 14195 Berlin, Germany
 Tel: +49-30-84131623
 Fax: +49-30-84131128
 Email: Poustka@mpg-berlin-dahlem.mpg.de
 This is the 5' sequence of the clone insert
 Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
 Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
 sequenced by DKFZ (German Cancer Research Center,
 Heidelberg/Germany) within the cDNA sequencing consortium of the
 German Genome Project.
 No SI sequence available.
 This clone (DKFZp686C2019) is available at the RZPD in Berlin.
 Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
 Berlin-Charlottenburg, GERMANY; Email: clone@zpd.de.

FEATURES

source

1. 550
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="DKFZp686C2019"
 /tissue_type="human skeletal muscle"
 /dev_stage="adult"
 /lab_host="DH10B"
 /clone_lib="686 (synonym: hlcc3)"
 /note="Vector: pTriplex2; Site 1: SfIra; Site 2: SfiIB;
 cDNA-collection"

BASE COUNT 135 a 159 c 161 g 95 t

ORIGIN

Alignment Scores:
 Pred. No.: 3.78e-15 Length: 550
 Score: 241.00 Matches: 57
 Percent Similarity: 65.66% Conservative: 8
 Best Local Similarity: 57.58% Mismatches: 16
 Query Match: 48.88% Indels: 18
 DB: 9 Gaps: 3

US-10-059-720-19 (1-97) x AL703189 (1-550)

QY 13 SerMetThrGlyGly-----GlnGlnMetGlyArgAspProAspLeuGluGlnArgAla 30
 Db 167 ACCATGACAGAGAGCCGAGCCAGCAGCATTTGGCAGG-----AGGGCAAGGTG 214
 QY 31 GluGluLeuAlaArgGluAsnGluGluLeuGluGluAlaGluGluGlnGlu 50
 Db 215 GAACAGTTATCTCCAGAGAAGAGAGAGAAAGAGATCCGAGGGAAGGATATAGATG 274
 QY 51 AsnAla-----GluLeuGluAlaGluThr 58
 Db 275 GCTCGACGCCAAATCCCGAACCGGAGGAGGTGACTGATACCTCCAGCGGAGACA 334
 QY 59 AspGlnLeuGluAspGluLysSerAlaLeuGlnThrGluLeuAlaAsnLeuLysGlu 78
 Db 335 GACCACTAGAGATGAGAGTCTGCTTTCGAGACCGAGATTGCCACCTGCTGAAGGAG 394

QY 79 LysGluLysLeuGluPheLeuAlaAlaHisArgProAlaCysLysIleProAsp 97
 Db 395 AAGGAAAACTAGAGTTTCATCTCTGGCAGCTCACCGACCTGCTGCAAGATCCCTGAT 451
 RESULT 4
 LOCUS BQ694927 551 bp mRNA linear EST 15-JUL-2002
 DEFINITION 1001070 Human Fat Cell 5'-Stretch Plus cDNA Library Homo sapiens
 cDNA 5', mRNA sequence.
 ACCESSION BQ694927
 VERSION BQ694927.1 GI:21820243
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 551)
 AUTHORS Yang, R.-Z., Shuldiner, A. and Gong, D.-W.
 TITLE EST analysis of human adipose gene expression
 JOURNAL Unpublished
 COMMENT Contact: Gong Da-Wei
 Division of Endocrinology, Diabetes and Nutrition
 University of Maryland
 660 Redwood St. HH497, Baltimore, MD 21201, USA
 Tel: 410 706 1672
 Fax: 410 706 1622
 Email: dgong@medicine.umaryland.edu
 PCR Primers
 FORWARD: CTCGGGAAGCGCGCCATTTGTGTGGT
 BACKWARD: AATACGACTACTATAGGCGGAATTGG
 Seq primer: GTTGGTACCGGGAATTC.
 FEATURES Location/Qualifiers
 1. 551
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /sex="Male and Female"
 /tissue_type="Adipose"
 /clone_lib="Human Fat Cell 5'-Stretch Plus cDNA Library"
 /note="Vector: lambdaTriplex"
 134 a 162 c 161 g 94 t
 BASE COUNT
 ORIGIN
 Alignment Scores:
 Pred. No.: 3.79e-15 Length: 551
 Score: 241.00 Matches: 57
 Percent Similarity: 65.66% Conservative: 8
 Best Local Similarity: 57.58% Mismatches: 16
 Query Match: 48.88% Indels: 18
 DB: 13 Gaps: 3
 US-10-059-720-19 (1-97) x BQ694927 (1-551)
 QY 13 SerMetThrGlyGly-----GlnGlnMetGlyArgAspProAspLeuGluGlnArgAla 30
 Db 190 ACCATGACAGAGAGCCGAGCCAGCAGCATTTGGCAGG-----AGGGCAAGGTG 237
 QY 31 GluGluLeuAlaArgGluAsnGluGluLeuGluGluAlaGluGluGlnGlu 50
 Db 238 GAACAGTTATCTCCAGAGAAGAGAGAAAGAGATCCGAGGGAAGGATATAGATG 297
 QY 51 AsnAla-----GluLeuGluAlaGluThr 58
 Db 298 GCTCGACGCCAAATCCCGAACCGGAGGAGGTGACTGATACCTCCAGCGGAGACA 357
 QY 59 AspGlnLeuGluAspGluLysSerAlaLeuGlnThrGluLeuAlaAsnLeuLysGlu 78
 Db 358 GACCACTAGAGATGAGAGTCTGCTTTCGAGACCGAGATTGCCACCTGCTGAAGGAG 417
 QY 79 LysGluLysLeuGluPheLeuAlaAlaHisArgProAlaCysLysIleProAsp 97
 Db 418 AAGGAAAACTAGAGTTTCATCTCTGGCAGCTCACCGACCTGCTGCAAGATCCCTGAT 474

```

RESULT 5
CA406906
LOCUS
DEFINITION 1003055 Human Fat Cell 5'-Stretch Plus cDNA Library Homo sapiens
CDNA 5', mRNA sequence.
ACCESSION CA406906
VERSION 1003055 Human Fat Cell 5'-Stretch Plus cDNA Library Homo sapiens
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE
AUTHORS Yang,R.-Z., Shuldiner,A. and Gong,D.-W.
TITLE EST analysis of human adipose gene expression
JOURNAL Unpublished
COMMENT Contact: Gong Da-Wei
Division of Endocrinology, Diabetes and Nutrition
University of Maryland
660 Redwood St, HH497, Baltimore, MD 21201, USA
Tel: 410 706 1672
Fax: 410 706 1622
Email: dgong@medicine.umaryland.edu
PCR Primers
FORWARD: CTCGGGAAGCGGCCATGTGTTGGT
BACKWARD: AATACGACTCATTATAGGCGGAATTGG
Seq primer: GTTGGTACCGGGAATTC.
FEATURES
source
1..565
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/sex="Male and Female"
/tissue type="Adipose"
/clone_lib="Human Fat Cell 5'-Stretch Plus cDNA Library"
/note="Vector: lambdaTriplex"
BASE COUNT 146 a 161 c 161 g 96 t 1 others
ORIGIN
Alignment Scores:
Pred. No.: 3 89e-15 Length: 565
Score: 241.00 Matches: 57
Percent Similarity: 65.66% Conservative: 8
Best Local Similarity: 57.58% Mismatches: 16
Query Match: 48.88% Indels: 18
DB: 14 Gaps: 3
US-10-059-720-19 (1-97) x CA406906 (1-565)
QY 13 SerMetThrGlyGly-----GlnGlnMetGlyArgAspProAspLeuGluGlnArgAla 30
Db 115 ACCATGACAGGAGGCGCGAGCGCAGCATTTGGCAGG-----AGGGCGAAGGTG 162
QY 31 GluGluLeuAlaArgGluAsnGluGluLeuGluGluGluAlaGluGluGluGluGlu 50
Db 163 GRACAGTTATCTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 222
QY 51 AsnAla-----GluLeuGluAlaGluThr 58
Db 223 GCTGCAGCCAAATGCCCAACCGAGGAGGAGCTGACTGATACATCTCCAGCGGAGACA 282
QY 59 AspGlnLeuGluAspGluLysSerAlaLeuGlnThrGluLeuAlaAsnLeuLeuLysGlu 78
Db 283 GACCACTAGAGATAGAGAGAGTCTGCTTTGCAGACGAGATTGCCAACCTGCTGAAGGAG 342
QY 79 LysGluLysLeuGluPheilleuAlaAlaHisArgProAlaCysIysIleProAsp 97
Db 343 AAGGAAAACTAGAGTTTCATCTTGGCGAGCTCACCGACCTGCTGCAAGATCCCTGAT 399
RESULT 6
BQ694699
LOCUS
DEFINITION 1000842 Human Fat Cell 5'-Stretch Plus cDNA Library Homo sapiens
CDNA 5', mRNA sequence.
ACCESSION BQ694699
VERSION 1000842 Human Fat Cell 5'-Stretch Plus cDNA Library Homo sapiens
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE
AUTHORS Yang,R.-Z., Shuldiner,A. and Gong,D.-W.
TITLE EST analysis of human adipose gene expression
JOURNAL Unpublished
COMMENT Contact: Gong Da-Wei
Division of Endocrinology, Diabetes and Nutrition
University of Maryland
660 Redwood St, HH497, Baltimore, MD 21201, USA
Tel: 410 706 1672
Fax: 410 706 1622
Email: dgong@medicine.umaryland.edu
PCR Primers
FORWARD: CTCGGGAAGCGGCCATGTGTTGGT
BACKWARD: AATACGACTCATTATAGGCGGAATTGG
Seq primer: GTTGGTACCGGGAATTC.
FEATURES
source
1..614
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/sex="Male and Female"
/tissue type="Adipose"
/clone_lib="Human Fat Cell 5'-Stretch Plus cDNA Library"
/note="Vector: lambdaTriplex"
BASE COUNT 145 a 188 c 172 g 109 t
ORIGIN
Alignment Scores:
Pred. No.: 4 22e-15 Length: 614
Score: 241.00 Matches: 57
Percent Similarity: 65.66% Conservative: 8
Best Local Similarity: 57.58% Mismatches: 16
Query Match: 48.88% Indels: 18
DB: 13 Gaps: 3
US-10-059-720-19 (1-97) x BQ694699 (1-614)
QY 13 SerMetThrGlyGly-----GlnGlnMetGlyArgAspProAspLeuGluGlnArgAla 30
Db 198 ACCATGACAGGAGGCGCGAGCGCAGCATTTGGCAGG-----AGGGCGAAGGTG 245
QY 31 GluGluLeuAlaArgGluAsnGluGluLeuGluLysGluAlaGluGluGluGluGlu 50
Db 246 GRACAGTTATCTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 305
QY 51 AsnAla-----GluLeuGluAlaGluThr 58
Db 306 GCTGCAGCCAAATGCCCAACCGAGGAGGAGCTGACTGATACATCTCCAGCGGAGACA 365
QY 59 AspGlnLeuGluAspGluLysSerAlaLeuGlnThrGluLeuAlaAsnLeuLeuLysGlu 78
Db 366 GACCACTAGAGATAGAGAGTCTGCTTTGCAGACGAGATTGCCAACCTGCTGAAGGAG 425
QY 79 LysGluLysLeuGluPheilleuAlaAlaHisArgProAlaCysIysIleProAsp 97
Db 426 AAGGAAAACTAGAGTTTCATCTTGGCGAGCTCACCGACCTGCTGCAAGATCCCTGAT 482
RESULT 7
BQ924297
LOCUS
DEFINITION HNC26-1-C12.R HNC (Human Normal Cartilage) Homo sapiens cdna, mRNA
sequence.
ACCESSION BQ924297
VERSION BQ924297.1 GI:14318820
KEYWORDS EST.

```

SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS	Kumar,S., Connor,J.R., Dodds,R.A., Halsey,W., Van Horn,M., Mao,J., Sathre,G., Mui,P., Agarwal,P., Badger,A.M., Lee,J.C., Gowen,M. and Lark,M.W.
TITLE	Identification and initial characterization of 5000 expressed sequenced tags (ESTs) each from adult human normal and osteocarthric cartilage cDNA libraries
JOURNAL	Osteoarthr. Cartil. 9 (7), 641-653 (2001)
MEDLINE	21482651
PUBMED	11597177
COMMENT	Contact: Sanjay Kumar UM2109 GlaxoSmithKline 709 Swedeland Road, P.O. Box 1539, King of Prussia, PA 19406, USA Tel: 610-270-7245 Fax: 610-270-5598 Email: sanjay.kumar-l@gsk.com Seq primer: T7:
FEATURES	Location/Qualifiers
source	1..615
/organism="Homo sapiens"	
/mol_type="mRNA"	
/db_xref="taxon:9606"	
/tissue_type="cartilage"	
/lab_host="E.coli DH10 B"	
/clone_lib="HNC (Human Normal Cartilage)"	
/notes=Vector: pSPORT I; Site_1: SalI; Site_2: NotI;	
Directional	
BASE COUNT	150 a 185 c 177 g 103 t
ORIGIN	
Alignment Scores:	
Pred. No.: 4.23e-15 Length: 615	
Score: 241.00 Matches: 57	
Percent Similarity: 65.66% Conservative: 8	
Best Local Similarity: 57.58% Mismatches: 16	
Query Match: 48.88% Indels: 18	
DB: 12 Gaps: 3	
US-10-059-720-19 (1-97) x BG924297 (1-615)	
QY 13 SerMetThrGlyGly-----GlnGlnMetGlyArgAspProaspLeuGluNArgAla 30	
Db 306 ACCATGACAGGGCGCGAGCGCAGCATTTGGCAGG-----AGGGCGCAAGGTG 353	
QY 31 GluGluLeuAlaArgGluAsnGluGluGluGluGluGluGluGluGluGluGluGlu 50	
Db 354 GAACAGTATTCTCCAGAT 413	
QY 51 AsnAla-----GluLeuGluAlaGluThr 58	
Db 414 GCTGCAGCCAAATGCCGACCGAGGAGGAGTGACTGATACATCTCCAGCGGAGACA 473	
QY 59 AspGlnLeuGluAspGluGluSerAlaLeuGlnThrGluLeuAlaAsnLeuGluGlu 78	
Db 474 GACCAACTAGAAGATGAGAAGTCGTCTTTCACAGCCGAGATTCGCCAACCCTGCTGAAGGAG 533	
QY 79 LysGluLysLeuGluPheIleLeuAlaAlaHisArgProAlaCysLysIleProasp 97	
Db 534 AAGGAAAACCTAGAGTTTCATCTCTGGCAGCTCACGACCTCGCTGCAGATCCCTGAT 590	
RESULT 8	
BQ694763	
LOCUS	BQ694763 Homo sapiens
DEFINITION	1000906 Human Fat Cell 5'-Stretch plus cDNA Library Homo sapiens
Accession	cDNA 5', mRNA sequence.
Version	BQ694763.1 GI:21820079
Keywords	EST.

Qy	59	AspGluLeuGluAspGluYSerAlaLeuGlnThrGluLeuAlaAsnLeuLeuYsGlu 78
Db	215	GACCAACTAGAGATGAGAGTCTGCTTTGACAGCCGAGATTGCCAACTGCTGAAGGAG 272
Qy	79	LysGluLeuGluPheIleLeuAlaAlaHisAtgProAlaCysIysIleProAsp 97
Db	275	AAGGAAAACTAGAGTTTCATCTCTGGCAGCTCACCGACCTGCTGCAAGATCCCTGAT 331
RESULT 12		
BI481329		667 bp mRNA linear EST 28-FEB-2000
LOCUS		H2RPE-1611 Human Retinal Pigment Epithelium (2) Homo sapiens CDNA
DEFINITION		5' similar to v-fos FBJ murine osteosarcoma viral oncogene homolog
ACCESSION	BI481329	1 GI:18999138
VERSION	EST.	
KEYWORDS	Homo sapiens (human)	
SOURCE	Homo sapiens	
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
REFERENCE	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	
AUTHORS	1 (bases 1 to 667)	
TITLE	Baraczynska,M., Mears,A.J., Zarepari,S., Farjo,R., Filipova,E.,	
JOURNAL	Yuan,Y., Macnee,S.P., Hughes,B. and Swaroop,A.	
COMMENT	Towards an expression profile of native human retinal pigment	
	epithelium: Identification of a non-redundant set of more than 110	
	genes	
	Unpublished	
	Contact: Swaroop, A.	
	Department of Ophthalmology and Visual Sciences	
	Kellogg Eye Center, University of Michigan	
	540 KEC, 1000 Wall St., Ann Arbor, MI 48105, USA	
	Tel: 734 615 2246	
	Fax: 734 647 0228	
	Email: swaroop@umich.edu	
	PCR Primers	
	FORWARD: M13/PUC-Reverse - cccagtcacagctgtgtataaacg	
	BACKWARD: M13/PUC-Forward - agcggataacaatttcacacagg	
	Seq primer: M13/PUC-Reverse.	
	Location/Qualifiers	
	1. .667	
	/organism="Homo sapiens"	
	/mol_type="mRNA"	
	/db_xref="taxon:9606"	
	/tissue_type="Native Retinal Pigment Epithelium sheets"	
	/dev_stage="juvenile"	
	/clone_lib="Human Retinal Pigment Epithelium (2)"	
	/note="Organ: Retina; Vector: pReport1"	
	BASE COUNT 159 a 199 c 189 g 117 t 3 others	
	ORIGIN	
	Alignment Scores:	
	Pred. No.: 4,59e-15 Length: 667	
	Score: 241.00 Matches: 57	
	Percent Similarity: 65.66% Conservative: 8	
	Best Local Similarity: 57.58% Mismatches: 16	
	Query Match: 48.88% Indels: 18	
	DB: 12 Gaps: 3	
	US-10-059-720-19 (1-97) x BI481329 (1-667)	
Qy	13	SerMetThrGlyGly-----GlnGlnMetGlyArgAspProAspLeuGlnArgAla 30
Db	165	ACCATGACAGGAGCGCGAGCGAGCATTTGGCAGG-----AGGGCGAAGGTG 212
Qy	31	GluGluLeuAlaArgGluAsnGluGluLeuGluLysGluAlaGluGluGlnGlu 50
Db	213	GACACGTATCTCCAGAGAGAGAGAGAAAGAGGAATCCGAGGAGGAAGATAGATG 272
Qy	51	AsnAla-----GluLeuGluAlaGluThr 58
Db	273	GCTGCGACCCAAATGCCCGAACCGGAGGAGGAGCTGACTGATACACTCCAGCGGAGACA 214

Qy	59	AspGlnLeuGluHisPheLeuSerAlaLeuGlnThrGluIleAlaAsnLeuLeuLysGlu	78
Db	333	GACCACCTAGAGATGAGAGTCTGCTTTTCAGACCCAGAGATGCCAACCTGCTGACGAG	392
Qy	79	LysGluLysLeuGluPheIleLeuAlaAlaHisArgProAlaCysLysIleProAsp	97
Db	393	AAGGAAANACTAGAGTTCTCTCTGGCAGCTCCACGACCTGCTGCAAGATCCCTGAT	449

RESULT 13
 BM686586
 LOCUS
 DEFINITION
 UI-E-CQ0-ado-c-12-0-UI.r1 UI-E-CQ0 Homo sapiens cDNA clone
 UI-E-CQ0-ado-c-12-0-UI 5', mRNA sequence.
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Homo sapiens (human)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 691)
 Bonaldo,M.F., Lennon,G. and Soares,M.B.
 Normalization and subtraction: two approaches to facilitate gene
 discovery
 Genome Res. 6 (9), 791-806 (1996)
 97044477
 8889548
 Contact: Soares, MB
 Coordinated Laboratory for Computational Genomics
 University of Iowa
 375 Newton Road, 4156 MBRF, Iowa City, IA 52242, USA
 Tel.: 319 335 8250
 Fax: 319 335 9565
 Email: bento-soares@uiowa.edu
 Tissue Procurement: Dr. Gregg Hageman
 cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
 cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
 Clone Distribution: Researchers may obtain clones from Research
 Genetics (www.resgen.com).
 Seq primer: M13 Reverse.

BASE COUNT	162 a	209 c	196 g	122 t	2 o
ORIGIN	System, supported by National Eye Institute				
Alignment Scores:					
Pred. No.:	4.74e-15			Length:	691
Score:	241.00			Matches:	57
Percent Similarity:	65.66%			Conservative:	58

Best Local Similarity:	57.58%	Mismatches:	16
Query Match:	48.88%	Indels:	18
DB:	12	Gaps:	3
US-10-059-720-19 (1-97) x BM686586 (1-691)			
QY	13	SerMetThrGlyGly-----GlGlnMetGlyArgAspProAspLeuGluGlnArgAla	30
Dd	197	ACCATGAGCAGAGGCCCGAGCCGACGAGCATTTGGCAGG-----AGGGGCACAGGTG	244
QY	31	GUUGlUeLuAlaArGGuAsnGUUGlUeLUgLUyGLuAlaGUUGlUeLUgLUgLUgLU	50
Dd	245	GAACAGTATTCTCCAGAAGAGAGAGAAAGGNGAATCCGAGGGGAAGGAATAAGATG	304
QY	51	AenAla-----GluLeuGluAlaGluThr	58
Dd	305	GTCTGAGCGCAAAATGCGCAACGGGAGGAGGAGTGACTGATACATCCAAGCGGAGACA	364
QY	59	ASPGLInLeuGluAspGluLYSerAlaLeuGlnThrGluileAlaasnLeuLeuLySGlu	78
Dd	365	GACCACACTAGAGATGAGAGAGTCGTTTCGACACCAGATTGCCAACCTGCTGAAGGAG	424
QY	79	LysGluLysLeuGIUpheIleLeuAlaAlaHisArgProAlaCysLysIleProasp	97
Dd	425	AAGGAAAAAATAGAGTTTCATCTCGCAGCTCACCGACCTGCTGCAAGATCCCTGAT	481
RESULT 14			
BG925313 713 bp mRNA linear EST 06-NOV-2000			
LOCUS HNC40-1-A5.R HNC (Human Normal Cartilage) Homo sapiens cDNA, mRNA			
DEFINITION sequence.			
ACCESSION BG925313			
VERSION BG925313.1 GI:14319836			
KEYWORDS EST.			
SOURCE Homo sapiens (human)			
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
REFERENCE Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.			
AUTHORS Kumar,S., Connor,J.R., Dadds,R.A., Halsley,W., Van Horn,M., Mao,J.,			
Sathe,G., Mui,P., Agarwal,P., Badger,A.M., Lee,J.C., Gowen,M. and			
Lark,M.W.			
TITLE Identification and initial characterization of 5000 expressed			
sequenced tags (ESTs) each from adult human normal and			
osteochondritic cartilage cDNA libraries			
JOURNAL Osteoarthr. Cartil. 9 (7), 641-653 (2001)			
MEDLINE 21482651			
PubMed 11597177			
COMMENT Contact: Sanjay Kumar			
UW2109			
GlaxoSmithKline			
709 Swedeland Road, P.O. Box 1539, King of Prussia, PA 19406, USA			
Tel: 610-270-7245			
Fax: 610-270-5598			
Email: sanjay.kumar-legsk.com			
Seq primer: 47:			
Location/Qualifiers			
1..713			
/organism="Homo sapiens"			
/mol_type="mRNA"			
/db_xref="taxon:9606"			
/tissue_type="cartilage"			
/lab_host="E.coli DH10 B"			
/clone_lib="HNC (Human Normal Cartilage)"			
/note=vector: pSPORT I; Site_1: SalI; Site_2: NotI;			
Directional"			
BASE COUNT	178 a	199 c	125 t
ORIGIN			1 others
Alignment Scores:			
Fred. No.:	4.89e-15	Length:	713
Score:	241.00	Matches:	57
Percent Similarity:	65.66%	Conservative:	8

BASE COUNT	178 a	199 c	210 g	125 t	1 others
ORIGIN	Directional*				
Alignment Scores:					
Pred. No.:	4.89e-15		Length: 713		
Score:	241.00		Matches: 57		
Percent Similarity:	65.66%		Conservative: 8		

Best Local Similarity:	57.58%	Mismatches:	16
Query Match:	48.88%	Indels:	18
DB:	12	Gaps:	3

US-10-059-720-19 (1-97) x BG925313 (1-713)

13	SerMetThrGlyGly	-----GlnClnMetClyArgAspProAspLeuGluInArgAla	30
	::::		::::
105	ACCATGACAGGGGGCCGACGACGATGTGGCAG	-----AGGGCCAGGTG	152
31	GluGluLeuAlaArgGluAsnGluClnLeuGluLysGluAlaGluGluLeuGluGlnGlu	50	
	::::	::::	
153	GAACAGTATTATCTCCAGAGAGAGAGAGAAAGGAGAAATCCGAGGGGAAAGAAATGATG	212	
51	AsnAla-	-----GluLeuGluAlaGluThr	58
213	GCTGACGCCAAATGCCGCAACCGGAGGGAGCTGACTGATACATCCCAAGCGGAGACA	272	
59	AspGlnLeuGluAspGluLysSerAlaLeuGlnThrGluIleAlaAsnLeuLeuLysGlu	78	
273	GAACCACTAGAGAGATGAGAAATCTGCTTTGCAGACCCGAGATTCACCACTGCTGAAGGAG	332	
79	LysGluLysLeuGluPheIleLeuAlaAlaHisArgProAlaCysIysIleProAsp	97	
333	AGAGAAACCTAGAGTTTATCTGGCACTCACCGACTGCTGCAAGATCCCTGAT	389	

RESULT 15	CA777910	755 bp	mRNA	linear	EST 03-DEC-2002
LOCUS	CA777910				
DEFINITION	ip15b02.x1 HR85 islet Homo sapiens cDNA clone IMAGE:521725 3' similar to SW:FOS_HUMAN P01100 P55-C-FOS PROTO-ONCOGENE PROTEIN ; mRNA sequence.				

NOTE: Site 2: XhoI: cDNA made by oligo-dT priming. Size selected on agarose gel. Average insert size ~1kb. 5' XhoI site was destroyed after directional cloning. Amplified once. Contact information: Hiroshi Inoue, MD, Metabolism Div. (Alan Permut Lab) Washington University School of Medicine, Box 8127, 660 South Euclid Ave. St. Louis, MO 63110, E-mail: hinoe@mgate.wustl.edu, Tel: 314-362-1915, Fax: 314-747-2692.

314-362-1916, FAX: 314-747-
162 a 248 c 209 g 136 t
BASE COUNT
ORIGIN

Alignment Scores:	5.18e-15	Length:	755
Pred. No.:	241.00	Matches:	57
Score:	65.66	Conservative:	8
Percent Similarity:	57.55%	Mismatches:	168
Best Local Similarity:	48.88%	Indels:	18
Query Match:	14	Gaps:	3
DB:			

US-10-059-720-19 (1-97) X CA777910 (1-755)

13	SerMetThrGlyGly	-----GlnGlnMetGlyArgAspProAspLeuGluGlnArgAla	30
			:::
			:::
402	ACCATGACAGGAGGCCGAGCCGACGATCTGGCAGG	-----AGGGCCCAAGAGTGG	449
			:::
			:::
31	GluGluLeuAlaAAGGluAsnGluGluLeuLysGluAlaGluGluLeuGluGlnGlu	50	
			:::
			:::
450	GAACTGTTATCTCCAGATGAGAGAGAGAAAAGGAGATCCGACGGGAAAGGCAATAGATG	509	
			:::
			:::
51	AsnAla	-----GluLeuGluAlaGluThr	58
			:::
			:::
510	GCTGCAGCAAAATGCTCCGCAACCGGAGGAGGAGCTGACTGATACACTCCAAAGCCGAGACA	569	
			:::
			:::
59	AspGlnLeuGluAsnGluLysSerAlaLeuGluThrGluLeuAlaAsnLeuLysGlu	78	
			:::
			:::
570	GACCAACTAGAAAGTAGAAGTCTGCTTCACAGCCGAGATGCCAACTCTGCTGAGGAG	629	
			:::
			:::
79	LysGluLysLeuGluPheIleLeuAlaAlaHisArgProAlaCysLysIleProAsp	97	
			:::
			:::
630	AAGGAAAACTAGAGTTATCTCTGCGAGCTCACCGACCTGCTGCAAGATCCCTGAT	686	
			:::
			:::

Search completed: November 16, 2003, 02:56:51
Job time : 1721 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: November 16, 2003, 02:28:12 ; Search time 273 Seconds

(without alignments)
1161.230 Million cell updates/sec

Title: US-10-059-720-19

Perfect score: 493

Sequence: 1 MDYKDDDKHMASMTGGQM.....EKEKLEFLAAHRAHPACKIPD 97

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2169961 seqs, 1634102185 residues

Total number of hits satisfying chosen parameters: 4339922

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-O=/cgn2_6/ptodata/1/pubpna/pct New PUB.seq:
-DB=Published Applications NA -QFMT=fastap -SUFFIX=rnpb -MINMATCH=0.1
-LOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blom62
-TRANS=puman40.cgi -LIST=45 -DOCLIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pct -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US10059720 -CGCN 1 1 347 -runat 14112003 185001_17978
-NCPU=6 -ICPU=3 -NO_WMAP -LARGQUERY -NEG_SCORES=0 -WAIT -DSFBJOCH=100
-LOGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-XGAPOP=6 -XGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA:

1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:
2: /cgn2_6/ptodata/1/pubpna/pct New PUB.seq:
3: /cgn2_6/ptodata/1/pubpna/US06_NEW PUB.seq:
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:
5: /cgn2_6/ptodata/1/pubpna/US07_NEW PUB.seq:
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:
7: /cgn2_6/ptodata/1/pubpna/US08_NEW PUB.seq:
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:
9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:
10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:
12: /cgn2_6/ptodata/1/pubpna/US09_NEW PUB.seq:
13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:
14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:
15: /cgn2_6/ptodata/1/pubpna/US10_NEW PUB.seq:
16: /cgn2_6/ptodata/1/pubpna/US60_NEW PUB.seq:
17: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB ID	Description
1	493	100.0	300	14	US-10-059-720-18

2	441	89.5	281	14	US-10-059-720-16	Sequence 16, Appl
3	350	71.0	324	14	US-10-059-720-10	Sequence 14, Appl
4	288	58.4	266	14	US-10-059-720-10	Sequence 10, Appl
5	245	49.7	403	10	US-09-980-352-11134	Sequence 11134, A
6	241	48.9	426	11	US-09-918-995-33976	Sequence 33976, A
7	241	48.9	2084	12	US-10-007-926A-318	Sequence 318, Appl
8	241	48.9	2135	12	US-09-971-392-8	Sequence 32, Appl
9	238	48.3	296	14	US-10-059-720-32	Sequence 7855, A
10	236	47.9	399	10	US-09-960-352-7855	Sequence 34766, A
11	235	47.7	435	11	US-09-918-995-34766	Sequence 12, Appl
12	229	46.5	262	14	US-10-059-720-12	Sequence 21268, A
13	228	46.2	451	11	US-09-918-995-27268	Sequence 16, Appl
14	216	43.8	273	14	US-10-070-676-16	Sequence 14906, A
15	216	43.8	415	10	US-09-960-352-14906	Sequence 38, Appl
16	210	42.6	1176	13	US-10-002-600-39	Sequence 1, Appl
17	206	41.8	435	10	US-09-785-298A-1	Sequence 25813, A
18	206	41.8	436	12	US-10-029-386-25813	Sequence 12113, A
19	206	41.8	600	12	US-10-029-386-12113	Sequence 20191, A
20	206	41.8	642	9	US-09-864-761-20191	Sequence 3420, Ap
21	206	41.8	1946	9	US-09-864-761-3420	Sequence 36, Appl
22	206	41.8	2227	13	US-10-002-600-36	Sequence 649, App
23	206	41.8	3565	12	US-09-873-319-649	Sequence 933, App
24	206	41.8	3565	12	US-09-960-706-993	Sequence 793, App
25	206	41.8	3565	12	US-09-873-367C-793	Sequence 15, Appl
26	206	41.8	3967	11	US-09-165-522-15	Sequence 201, Appl
27	206	41.8	6210	12	US-10-101-510-201	Sequence 10656, A
28	204	41.4	445	11	US-09-918-995-10656	Sequence 24355, Ap
29	192	38.9	497	14	US-10-066-543-2435	Sequence 34550, A
30	189	38.3	429	11	US-09-918-995-34550	Sequence 436, App
31	186	37.7	300	10	US-09-920-300A-436	Sequence 436, App
32	186	37.7	300	12	US-10-099-926-436	Sequence 436, App
33	186	37.7	300	13	US-10-033-528-436	Sequence 436, App
34	180.5	36.6	1017	12	US-10-004-113-18	Sequence 18, Appl
35	180.5	36.6	3775	10	US-09-954-456-2210	Sequence 2210, Ap
36	180.5	36.6	3775	10	US-09-880-107-2221	Sequence 404, App
37	180.5	36.6	3775	12	US-09-873-319-404	Sequence 649, App
38	180.5	36.6	3775	12	US-09-960-706-649	Sequence 698, App
39	180.5	36.6	3775	12	US-09-873-367C-698	Sequence 60, Appl
40	180.5	36.6	3775	12	US-10-210-120-60	Sequence 17, Appl
41	180.5	36.6	3775	12	US-10-004-113-17	Sequence 133, App
42	180.5	36.6	3775	14	US-10-205-843-133	Sequence 31, Appl
43	180.5	36.6	3851	12	US-10-247-671-31	Sequence 12898, A
44	180.5	36.6	3876	14	US-10-198-846-12898	Sequence 8, Appl
45	173.5	35.2	230	14	US-10-059-720-8	

ALIGNMENTS

RESULT 1
US-10-059-720-18
; Sequence 18, Application US/10059720
; Publication No. US20030027314A1
; GENERAL INFORMATION:

APPLICANT: VINSON, Charles R.

KEYLOV, Dmitry

TITLE OF INVENTION: EXTENSION OF A PROTEIN-PROTEIN INTERACTION SURFACE TO INACTIVATE THE FUNCTION OF A CELLULAR PROTEIN

NUMBER OF SEQUENCES: 64

CORRESPONDENCE ADDRESS:

ADDRESSEE: MORGAN & FINNEGAN, L.L.P.

STREET: 345 Park Avenue

CITY: New York

STATE: NY

COUNTRY: USA

ZIP: 10154-0053

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: MS WORD 97

SOFTWARE: ASCII

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/10/059,720

```

; FILING DATE: 29-Jan-2002
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/001,654
; FILING DATE: 31-JUL-1995
; APPLICATION NUMBER: 60/018,496
; FILING DATE: 29-MAY-1996
; APPLICATION NUMBER: 08/690,011
; FILING DATE: 31-JULY-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Serunian, Leslie A.
; REGISTRATION NUMBER: 35,353
; REFERENCE/DOCKET NUMBER: 2026-4199052
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212)751-6849
; TELEFAX: (212)751-6849
; INFORMATION FOR SEQ ID NO: 18:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 300 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: unknown
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; SEQUENCE DESCRIPTION: SEQ ID NO: 18:
US-10-059-720-18
Alignment Scores:
Pred. No.: 4,34e-52 Length: 300
Score: 493.00 Matches: 97
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 14 Gaps: 0
US-10-059-720-19 (1-97) x US-10-059-720-18 (1-300)
QY 1 MetAspTyrIysAspAspAspLysHisMetAlaSerMetThrGlyGlnGlnMet 20
Db 3 ATGCGACTACAGGACGACGATGACATGCTAGCATGCTGCTGAGCAAAATG 62
QY 21 GlyArgAspProAspLeuGlnArgAlaGluLeuAlaGluLeuAlaGluLeu 40
Db 63 GGTGGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 122
QY 41 GluLysGluAlaGluLeuGluGlnGluAsnAlaGluLeuGluAlaGluThrAspGln 60
Db 123 GAAAGAGGCGGAGAGCTGGAGCAGGAAAGCTGAACTCGAGCGGAGACACCA 182
QY 61 LeuGluAspGluLysSerAlaLeuGlnThrGluAlaAlaAsnLeuLysGlu 80
Db 183 CTAGAAGATGAGAGTCTGCTTTGACAGCCGAGATTGCCAAGCTGCTGAAGGAGGAA 242
QY 81 LysLeuGluPheIleLeuAlaHisArgProAlaCysLysIleProAsp 97
Db 243 AACCTAGATTATCTCTGCGACTACCGACTGCTGCTGCAAGATCCCTGAT 293
RESULT 2
US-10-059-720-16
; Sequence 16, Application US/10059720
; Publication No. US20030027314A1
; GENERAL INFORMATION:
; APPLICANT: VINSON, Charles R.
; KEYLOV, Dmitry
; TITLE OF INVENTION: EXTENSION OF A PROTEIN-PROTEIN
; INTERACTION SURFACE TO INACTIVATE THE FUNCTION OF A
; CELLULAR PROTEIN
; NUMBER OF SEQUENCES: 64
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORGAN & FINNEGAN, L.L.P.
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10154-0053
; FILING DATE: 29-Jan-2002
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/001,654
; FILING DATE: 31-JUL-1995
; APPLICATION NUMBER: 60/018,496
; FILING DATE: 29-MAY-1996
; APPLICATION NUMBER: 08/690,011
; FILING DATE: 31-JULY-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Serunian, Leslie A.
; REGISTRATION NUMBER: 35,353
; REFERENCE/DOCKET NUMBER: 2026-4199052
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212)751-6849
; TELEFAX: (212)751-6849
; INFORMATION FOR SEQ ID NO: 18:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 300 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: unknown
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; SEQUENCE DESCRIPTION: SEQ ID NO: 18:
US-10-059-720-18
Alignment Scores:
Pred. No.: 4,34e-52 Length: 300
Score: 493.00 Matches: 97
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 14 Gaps: 0
US-10-059-720-19 (1-97) x US-10-059-720-16 (1-281)
QY 10 HisMetAlaSerMetThrGlyGlnGlnMetGlyArgAspProAspLeuGluGlnArg 29
Db 6 CATATGCTGATGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 65
QY 30 AlaGluGluLeuAlaArgGluAsnGluLeuGluLysGluAlaGluGluGln 49
Db 66 GCTGAGGAACCTGGCCCGTGAAAGAGAGCTGGAAGAGAGAGAGAGAGAGAG 125
QY 50 GluAsnAlaGluLeuGluAlaGluThrAspGlnLeuGluAspGluLysSerAlaLeuGln 69
Db 126 GAAACGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 185
QY 70 ThrGluIleAlaAsnLeuLysGluLysGluLysGluLysGluPheIleLeuAlaHis 89
Db 186 ACCGAGATTGCCAACCCTGCTGAGGAGAGAGAGAGAGAGAGAGAGAGAGAG 245
QY 90 ArgProAlaCysLysIleProAsp 97
Db 246 CGACCTGCTGCAAGATCCCTGAT 269
RESULT 3
US-10-059-720-14
; Sequence 14, Application US/10059720
; Publication No. US20030027314A1
; GENERAL INFORMATION:
; APPLICANT: VINSON, Charles R.
; KEYLOV, Dmitry
; TITLE OF INVENTION: EXTENSION OF A PROTEIN-PROTEIN
; INTERACTION SURFACE TO INACTIVATE THE FUNCTION OF A
; CELLULAR PROTEIN
; NUMBER OF SEQUENCES: 64
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: MS WORD 97
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/059,720
; FILING DATE: 29-Jan-2002
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/001,654
; FILING DATE: 31-JUL-1995
; APPLICATION NUMBER: 60/018,496
; FILING DATE: 29-MAY-1996
; APPLICATION NUMBER: 08/690,011
; FILING DATE: 31-JULY-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Serunian, Leslie A.
; REGISTRATION NUMBER: 35,353
; REFERENCE/DOCKET NUMBER: 2026-4199052
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212)751-6849
; TELEFAX: (212)751-6849
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 281 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: unknown
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; SEQUENCE DESCRIPTION: SEQ ID NO: 16:
US-10-059-720-16
Alignment Scores:
Pred. No.: 1,01e-45 Length: 281
Score: 441.00 Matches: 88
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 89.45% Indels: 0
DB: 14 Gaps: 0
US-10-059-720-19 (1-97) x US-10-059-720-16 (1-281)
QY 10 HisMetAlaSerMetThrGlyGlnGlnMetGlyArgAspProAspLeuGluGlnArg 29
Db 6 CATATGCTGATGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 65
QY 30 AlaGluGluLeuAlaArgGluAsnGluLeuGluLysGluAlaGluGluGln 49
Db 66 GCTGAGGAACCTGGCCCGTGAAAGAGAGCTGGAAGAGAGAGAGAGAGAGAG 125
QY 50 GluAsnAlaGluLeuGluAlaGluThrAspGlnLeuGluAspGluLysSerAlaLeuGln 69
Db 126 GAAACGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 185
QY 70 ThrGluIleAlaAsnLeuLysGluLysGluLysGluLysGluPheIleLeuAlaHis 89
Db 186 ACCGAGATTGCCAACCCTGCTGAGGAGAGAGAGAGAGAGAGAGAGAGAGAG 245
QY 90 ArgProAlaCysLysIleProAsp 97
Db 246 CGACCTGCTGCAAGATCCCTGAT 269
RESULT 3
US-10-059-720-14
; Sequence 14, Application US/10059720
; Publication No. US20030027314A1
; GENERAL INFORMATION:
; APPLICANT: VINSON, Charles R.
; KEYLOV, Dmitry
; TITLE OF INVENTION: EXTENSION OF A PROTEIN-PROTEIN
; INTERACTION SURFACE TO INACTIVATE THE FUNCTION OF A
; CELLULAR PROTEIN
; NUMBER OF SEQUENCES: 64
```

;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: MORGAN & FINNEGAN, L.L.P.
;; STREET: 345 Park Avenue
;; CITY: New York
;; STATE: NY
;; COUNTRY: USA
;; ZIP: 10154-0053
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: MS WORD 97
;; SOFTWARE: ASCII
;; CURRENT APPLICATION DATA:
;; FILING DATE: 29-Jan-2002
;; APPLICATION NUMBER: US/10/059,720
;; PRIOR APPLICATION DATA:
;; FILING DATE: 29-Jan-2002
;; APPLICATION NUMBER: 60/001,654
;; FILING DATE: 31-JUL-1995
;; APPLICATION NUMBER: 60/018,496
;; FILING DATE: 29-MAY-1996
;; APPLICATION NUMBER: 08/690,011
;; FILING DATE: 31-JULY-1996
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Serunian, Leslie A.
;; REGISTRATION NUMBER: 35,353
;; REFERENCE/DOCKET NUMBER: 2026-4199US2
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (212)751-6849
;; INFORMATION FOR SEQ ID NO: 14:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 324 base pairs
;; TYPE: nucleic acid
;; STRANDEDNESS: unknown
;; TOPOLOGY: linear
;; MOLECULE TYPE: DNA (genomic)
;; SEQUENCE DESCRIPTION: SEQ ID NO: 14:
US-10-059-720-14
Alignment Scores:
Pred. No.: 1,87e-34 Length: 324
Score: 350.00 Matches: 72
Percent Similarity: 74.53% Conservative: 7
Best Local Similarity: 67.92% Mismatches: 17
Query Match: 70.99% Indels: 10
DB: 14 Gaps: 2
US-10-059-720-19 (1-97) x US-10-059-720-14 (1-324)
QY 1 MetAspTylLysAspAspAspLysHisMetAlaSerMetThrGlyGlyGlnMet 20
DB 3 ATGGACTACAGGACGACGATGACAGCATATGGCTAGCATATGGCTGGACAGCAATG 62
QY 21 GlyArgAspProAspLeuGluGlnArgAlaGluGluLeuAlaArgGluAsnGluLeu 40
DB 63 GTCTGGGATCCCAAGGTGGACAGTATCTCCAGAA---GAAGAAGAGAAAAGGAGATC 119
QY 41 GluLysGluAlaGluGluGlnGluGlnGluAlaGlu----- 53
DB 120 CGAAGGAAAGGATTAAGATGGCTGAGCAATATCCGCAACCGAGGAGGAGTACT 179
QY 54 -----LeuGluAlaGluThrAspGlnLeuGluAspGluLysSerAlaLeuGlnThrGlu 71
DB 180 GATACACTCCAGCGGAGACAGACCACTAGAGATGAGAGTCTGCTTTGACAGCCGAG 239
QY 72 IleAlaAsnLeuLeuLysGluLysGluLysLeuGluPheIleuAlaAlaHisArgPro 91
DB 240 ATTGCCAACCTGCTGAAGAGAGAGGAAAAAATAGAGTTTCCTTGGCAGCTCACCGRCT 299
QY 92 AlaCysLysIleProAsp 97
DB 300 GCCTGCAAGATCCCTGAT 317

RESULT 4
US-10-059-720-10
; Sequence 10, Application US/10059720
; Publication No. US20030027314A1
; GENERAL INFORMATION:
; APPLICANT: VINSON, Charles R.
; KRYLOV, Dmitry
; TITLE OF INVENTION: EXTENSION OF A PROTEIN-PROTEIN
; INTERACTION SURFACE TO INACTIVATE THE FUNCTION OF A
; CELLULAR PROTEIN
; NUMBER OF SEQUENCES: 64
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORGAN & FINNEGAN, L.L.P.
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10154-0053
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: MS WORD 97
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/059,720
; FILING DATE: 29-Jan-2002
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/001,654
; FILING DATE: 31-JUL-1995
; APPLICATION NUMBER: 60/018,496
; FILING DATE: 29-MAY-1996
; APPLICATION NUMBER: 08/690,011
; FILING DATE: 31-JULY-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Serunian, Leslie A.
; REGISTRATION NUMBER: 35,353
; REFERENCE/DOCKET NUMBER: 2026-4199US2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212)751-6849
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 266 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: unknown
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; SEQUENCE DESCRIPTION: SEQ ID NO: 10:
US-10-059-720-10
Alignment Scores:
Pred. No.: 6.3e-27 Length: 266
Score: 288.00 Matches: 59
Percent Similarity: 82.67% Conservative: 13
Best Local Similarity: 78.67% Mismatches: 3
Query Match: 58.42% Indels: 0
DB: 14 Gaps: 0
US-10-059-720-19 (1-97) x US-10-059-720-10 (1-266)
QY 1 MetAspTylLysAspAspAspLysHisMetAlaSerMetThrGlyGlyGlnMet 20
DB 3 ATGGACTACAGGACGACGATGACAGCATATGGCTAGCATATGGCTGGACAGCAATG 62
QY 21 GlyArgAspProAspLeuGluGlnArgAlaGluGluLeuAlaArgGluAsnGluLeu 40
DB 63 GTCTGGGATCCCTGACCTGGAAACACGCTGTGAGGAACCTGCCCGGAAACCGAGAGCTG 122
QY 41 GluLysGluAlaGluGluLeuGluGlnGlnAlaGluLeuGluAlaGluThrAspGln 60
DB 123 GAAAAAGAGGCCGAGAGAGCTGGAGAGGAACTGGCAGAACTCGAGAACAGAGTGGCAGTG 182
QY 61 LeuGluAspGluLysSerAlaLeuGlnThrGluIleAlaAsnLeu 75


```
QY 49 GlnGluAenAlaGluLeuGluAlaGluThrAspGlnLeuGluAspGluLysSerAlaLeu 68
Db 121 CAGGAAACGCTGAACCTCGACAGCAAGTGTGGAGTTGGAAGTCTAAATGACCGCTG 180
QY 69 GlnThrGluLeuAlaAsnLeuLysGluLysGluLysLeuGluPheLeu 85
Db 181 CGCAGGAAGTGGAACTGAGCTGGAGCTGAACCTGACACGCTGCGGGGTATC 231

RESULT 10
US-09-960-352-7855
; Sequence 7855, Application US/09960352
; Patent No. US20020137139A1
; GENERAL INFORMATION:
; APPLICANT: Warren, Wesley C.
; APPLICANT: Tao, Mengbing
; APPLICANT: Byatt, John C.
; APPLICANT: Mathialagan, Nagappan
; TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
; FILE REFERENCE: 16511.006/37-21(10298)C
; CURRENT APPLICATION NUMBER: US/09/960.352
; CURRENT FILING DATE: 2001-09-24
; NUMBER OF SEQ ID NOS: 15112
; SEQ ID NO 7855
; LENGTH: 399
; TYPE: DNA
; ORGANISM: Bos taurus
; OTHER INFORMATION: Clone ID: 34-LIB188-026-Q1-EI-A10
US-09-960-352-7855

Alignment Scores:
Pred. No.: 2,54e-20 Length: 399
Score: 236.00 Matches: 57
Percent Similarity: 64.36% Conservative: 8
Best Local Similarity: 56.44% Mismatches: 18
Query Match: 47.87% Indels: 18
Gaps: 10

US-10-059-720-19 (1-97) x US-09-960-352-7855 (1-399)
QY 11 MetAlaSerMetThrGlyGly-----GlnGlnMetGlyArgAspProAspLeuGluGln 28
Db 104 ATGAACCATGACAGAGGAGGAGCTCAGAGCATTTGCCGG-----AGGGGC 151
QY 29 ArgAlaGluLeuAlaArgGluAsnGluLeuGluLysGluAlaGluGluLeuGlu 48
Db 152 AAGGTAGAACAGTGTGTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 211
QY 49 GlnGluAsnAla-----GluLeuGluAla 56
Db 212 AAGATGGCTGACGCCAAATGCCGAAACCGGAGGAGGAGCTGACTGACACCTCCAGCG 271
QY 57 GluThrAspGlnLeuGluAspGluLysSerAlaLeuGlnThrGluLeuAlaAsnLeu 76
Db 272 GAGACAGCCAACTAGATAATGAGAGTTCGCTTTGACAGACAGAGATTCCTGCTG 331
QY 77 LysGluLysGluLysLeuGluPheLeuAlaAlaHisArgProAlaCysLysIlePro 96
Db 332 AAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 391
QY 97 Asp 97
Db 392 GAT 394

RESULT 11
US-09-918-995-34766
; Sequence 34766, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; FROM VARIOUS CDNA LIBRARIES
; FILE REFERENCE: 20411-756
```

```
; CURRENT APPLICATION NUMBER: US/09/918,995
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US/09/235,076
; PRIOR FILING DATE: 1999-01-20
; NUMBER OF SEQ ID NOS: 38054
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 34766
; LENGTH: 435
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-918-995-34766

Alignment Scores:
Pred. No.: 3,72e-20 Length: 435
Score: 235.00 Matches: 56
Percent Similarity: 64.65% Conservative: 8
Best Local Similarity: 56.57% Mismatches: 17
Query Match: 47.67% Indels: 18
Gaps: 3

US-10-059-720-19 (1-97) x US-09-918-995-34766 (1-435)
QY 13 SerMetThrGlyGly-----GlnGlnMetGlyArgAspProAspLeuGluGlnArgAla 30
Db 12 ACCATGACAGAGGCGGAGCGCAGACGATTCGCGAG-----AGGGCGAAGGTG 59
QY 31 GluGluLeuAlaArgGluAsnGluGluLeuGluLysGluAlaGluGluLeuGluGln 50
Db 60 GAACAGTTATCTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGATG 119
QY 51 AsnAla-----GluLeuGluAlaGluThr 58
Db 120 GCTGCGACCAATGCCGCAACCGGAGGAGGAGCTGACTGATACCTCCAGCGGAGACA 179
QY 59 AspGlnLeuGluAspGluLysSerAlaLeuGlnThrGluLeuAlaAsnLeuLysGlu 78
Db 180 GACCAACTAGAGATGAGAGTCTGCTTTGACAGCGGAGATTCGCAACCTGCTGAGGAG 239
QY 79 LysGluLysGluPheLeuAlaAlaHisArgProAlaCysLysIleProAsp 97
Db 240 AAGGAAACCTAGAGATGATCTCGGAGCTCACCAGCTCCCTGCGAAGATCCCTGAT 296

RESULT 12
US-10-059-720-12
; Sequence 12, Application US/10059720
; Publication No. US20030027314A1
; GENERAL INFORMATION:
; APPLICANT: VINSON, Charles R.
; APPLICANT: KRYLOV, Dmitry
; TITLE OF INVENTION: EXTENSION OF A PROTEIN-PROTEIN
; INTERACTION SURFACE TO INACTIVATE THE FUNCTION OF A
; CELLULAR PROTEIN
; NUMBER OF SEQUENCES: 64
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORGAN & FINNEGAN, L.L.P.
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10154-0053
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: MS WORD 97
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/059,720
; FILING DATE: 29-Jan-2002
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/001,654
; FILING DATE: 31-JUL-1995
; APPLICATION NUMBER: 60/018,496
; FILING DATE: 29-MAY-1996
```



```
; APPLICATION NUMBER: 08/690,011
; FILING DATE: 31-JULY-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Serunian, Leslie A.
; REGISTRATION NUMBER: 35,353
; REFERENCE/DOCKET NUMBER: 2026-4199US2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212)758-4800
; TELEFAX: (212)751-6849
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 262 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: unknown
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; SEQUENCE DESCRIPTION: SEQ ID NO: 12:
US-10-059-720-12
Alignment Scores:
Pred. No.: 1.12e-19 Length: 262
Score: 229.00 Matches: 50
Percent Similarity: 67.86% Conservative: 7
Best Local Similarity: 59.52% Mismatches: 17
Query Match: 46.45% Indels: 10
DB: 14 Gaps: 2
US-10-059-720-19 (1-97) x US-10-059-720-12 (1-262)
QY 23 AspProAspLeuGluGlnArgAlaGluLeuAlaArgGluAsnGluGluLeuGluLys 42
Db 2 GATCCCAAGGTGGAACAGTATCTCCAGAA--GAAGAGAGAAAGAGGATCCGAGG 58
QY 43 GluAlaGluLeuGluGlnGluAsnAlaGlu----- 53
Db 59 GAAAGAAATAGATGCTGCGACCAATGCCGCAACCGGAGGAGGAGTGAATACA 118
QY 54 LeuGluAlaGluThrAspGlnLeuGluAspGluLysSerAlaLeuGlnThrGluLeuAla 73
Db 119 CTCCAAGCGAGACACCACTAGAGATGAGAGTCTGCTTGCAGACCGAGATTGCC 178
QY 74 AsnLeuLeuLysGluLysGluLysLeuGluPheIleLeuAlaAlaHisArgProAlaCys 93
Db 179 AACCTGCTGAAGGAGAGAGAAACTAGAGTTTCATCTTGCAGCTCACCAGCTGCTGC 238
QY 94 LysIleProAsp 97
Db 239 AAGATCCCTGAT 250
RESULT 13
US-09-918-995-27268
; Sequence 27268, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; FILE REFERENCE: 20411-756
; CURRENT APPLICATION NUMBER: US/09/918,995
; PRIOR FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US/09/235,076
; NUMBER OF SEQ ID NOS: 38054
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 27268
; LENGTH: 451
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)-(451)
; OTHER INFORMATION: n = A,T,C or G
US-09-918-995-27268
; APPLICATION NUMBER: 08/690,011
; FILING DATE: 31-JULY-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Serunian, Leslie A.
; REGISTRATION NUMBER: 35,353
; REFERENCE/DOCKET NUMBER: 2026-4199US2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212)758-4800
; TELEFAX: (212)751-6849
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 262 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: unknown
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; SEQUENCE DESCRIPTION: SEQ ID NO: 12:
US-10-059-720-12
Alignment Scores:
Pred. No.: 1.12e-19 Length: 262
Score: 229.00 Matches: 50
Percent Similarity: 67.86% Conservative: 7
Best Local Similarity: 59.52% Mismatches: 17
Query Match: 46.45% Indels: 10
DB: 14 Gaps: 2
US-10-059-720-19 (1-97) x US-10-059-720-12 (1-262)
QY 23 AspProAspLeuGluGlnArgAlaGluLeuAlaArgGluAsnGluGluLeuGluLys 42
Db 2 GATCCCAAGGTGGAACAGTATCTCCAGAA--GAAGAGAGAAAGAGGATCCGAGG 58
QY 43 GluAlaGluLeuGluGlnGluAsnAlaGlu----- 53
Db 59 GAAAGAAATAGATGCTGCGACCAATGCCGCAACCGGAGGAGGAGTGAATACA 118
QY 54 LeuGluAlaGluThrAspGlnLeuGluAspGluLysSerAlaLeuGlnThrGluLeuAla 73
Db 119 CTCCAAGCGAGACACCACTAGAGATGAGAGTCTGCTTGCAGACCGAGATTGCC 178
QY 74 AsnLeuLeuLysGluLysGluLysLeuGluPheIleLeuAlaAlaHisArgProAlaCys 93
Db 179 AACCTGCTGAAGGAGAGAGAAACTAGAGTTTCATCTTGCAGCTCACCAGCTGCTGC 238
QY 94 LysIleProAsp 97
Db 239 AAGATCCCTGAT 250
RESULT 14
US-10-070-676-16/c
; Sequence 16, Application US/10070676
; Publication No. US20030059788A1
; GENERAL INFORMATION:
; APPLICANT: Toque, Bruno
; APPLICANT: Bracco, Laurent
; APPLICANT: Schweighoffer, Fabien
; TITLE OF INVENTION: Genetic Markers of Toxicity, Preparation
; FILE REFERENCE: 50146/003002
; CURRENT APPLICATION NUMBER: US/10/070,676
; PRIOR FILING DATE: 2002-03-06
; PRIOR APPLICATION NUMBER: PCT/FR00/02503
; PRIOR FILING DATE: 2000-09-12
; PRIOR APPLICATION NUMBER: FR 99/11405
; PRIOR FILING DATE: 1999-09-13
; PRIOR APPLICATION NUMBER: 09/456,370
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16
; LENGTH: 273
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-10-070-676-16
Alignment Scores:
Pred. No.: 4.69e-18 Length: 273
Score: 216.00 Matches: 43
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 97.73% Mismatches: 0
Query Match: 43.81% Indels: 0
DB: 14 Gaps: 0
US-10-059-720-19 (1-97) x US-10-070-676-16 (1-273)
QY 54 LeuGluAlaGluThrAspGlnLeuGluAspGluLysSerAlaLeuGlnThrGluLeuAla 73
Db 236 CTCCAAGCGGAGACAGACCACTAGAGATGAGAGTCTGCTTGCAGACCGAGATTGCC 177
QY 74 AsnLeuLeuLysGluLysGluLysLeuGluPheIleLeuAlaAlaHisArgProAlaCys 93
```

```
Db 176 AACCTGCTGAGGAGAGAGAACTAGAGTTTCCTCTGGCAGCTCACCGACCTGCCTGC 117
Qy 94 LysileProAsp 97
Db 116 AAGATCCCTGAT 105

RESULT 15
US-09-960-352-14906
; Sequence 14906, Application US/09960352
; Patent No. US20020137139A1
; GENERAL INFORMATION:
; APPLICANT: Warren, Wesley C.
; APPLICANT: Tao, Nengbing
; APPLICANT: Byatt, John C.
; APPLICANT: Mathialagan, Nagappan
; TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
; TITLE OF INVENTION: MUSCLE AND FAT DEPOSITION
; FILE REFERENCE: 16511.006/37-21(10298)C
; CURRENT APPLICATION NUMBER: US/09/960.352
; CURRENT FILING DATE: 2001-09-24
; NUMBER OF SEQ ID NOS: 15112
; SEQ ID NO 14906
; LENGTH: 415
; TYPE: DNA
; ORGANISM: Bos taurus
; OTHER INFORMATION: Clone ID: 64-LIB188-010-Q1-E1-H12
US-09-960-352-14906

Alignment Scores:
Pred. No.: 7.67e-18 Length: 415
Score: 216.00 Matches: 52
Percent Similarity: 63.37% Conservative: 12
Best Local Similarity: 51.49% Mismatches: 19
Query Match: 43.81% Indels: 18
DB: 10 Gaps: 3

US-10-059-720-19 (1-97) x US-09-960-352-14906 (1-415)
Qy 11 MetAlaSerMetThrGlyGly-----GlnGlnMetGlyArgAspProAspLeuGluGln 28
Db 120 ATGAAACACCATGACAGGAGGCACAGCTCAGAGCATTTGCCGG-----AAAGGC 167
Qy 29 ArgAlaGluGluLeuAlaArgGluAsnGluLeuGluLysGluAlaGluGluLeuGlu 48
Db 168 AAGGTAGAACAGTTGTCTCCCAAGAGAGAGAGAAAGGCGGACTCCGAATGGAACGGAAT 227
Qy 49 GlnGluAsnAla-----GluLeuGluAla 56
Db 228 AGGATGGCTGCAGCCAAATGCCGAACCGGAGGAGGAGTACTGACACCTCCAAATCG 287
Qy 57 GluThrAspGlnLeuGluAspGluLysSerAlaLeuGlnThrGluIleAlaAsnLeuLeu 76
Db 288 GAGACAGACCACTTACAAATATGAGAGAGTCACTTTGCAGACAGAGATTGCCAATCTGCTG 347
Qy 77 LysGluLysGluLysLeuGluPheIleLeuAlaAlaHisArgProAlaCysLysIlePro 96
Db 348 AAGGAGATGAAATAACTCGATTTCCTTAGCGGCTCATCGACCTGCTGCAAGATCCC 407
Qy 97 Asp 97
Db 408 GAT 410
```

Search completed: November 16, 2003, 03:46:31
Job time : 275 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: November 16, 2003, 01:32:57 ; Search time 56 Seconds
(without alignments)
764.539 Million cell updates/sec

Title: US-10-059-720-19

Perfect score: 493

Sequence: 1 MDYKDDDDKHWASWTGQQM.....EKEKLEFILAHRPACKIPD 97

Scoring table:

BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame_p2n.model -DEV=xlh
-O=/cgn2_1/uspto/spool/US10059720/runat 14112003 185000 17910/app query.fasta_1.263
-DB=Issued_Patents_NA -QWTF=fastap -SUFFIX=rml -MINMATCH=0.1 -LOOPEL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pco -NORM=ext -HEAPSIZE=500 -MNLLEN=0 -MAXLEN=2000000000
-USER=US10059720 @cgn 1.1.56 @runat 14112003 185000 17910 -NCPUS=6 -ICPU=3
-NO MWAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents_NA:

1: /cgn2_6/ptodata/1/ina/5A.COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B.COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A.COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B.COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PTUS.COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	493	100.0	294	2	US-08-690-011A-18
2	493	100.0	300	4	US-09-299-495F-18
3	441	89.5	281	4	US-09-299-495F-16
4	436	88.4	281	2	US-08-690-011A-16
5	350	71.0	324	4	US-09-299-495F-14
6	349	70.8	318	2	US-08-690-011A-14
7	288	58.4	266	4	US-08-690-011A-10
8	288	58.4	266	4	US-09-299-495F-10
9	238	48.3	296	2	US-08-690-011A-32
10	238	48.3	296	4	US-09-299-495F-32
11	229	46.5	262	2	US-08-690-011A-12
12	229	46.5	262	4	US-09-299-495F-12

13	206	41.8	435	4	US-09-765-298A-1	Sequence 1, Appli
14	206	41.8	3565	1	US-08-306-691B-44	Sequence 44, Appl
15	206	41.8	3565	5	PCT-US93-06251-20	Sequence 20, Appl
16	173.5	35.2	288	2	US-08-690-011A-8	Sequence 8, Appli
17	173.5	35.2	290	4	US-09-299-495F-8	Sequence 8, Appli
18	173	35.1	4144	1	US-08-218-686-1	Sequence 1, Appli
19	173	35.1	4144	3	US-08-460-242-1	Sequence 1, Appli
20	172	34.9	353	4	US-09-299-495F-20	Sequence 20, Appl
21	171.5	34.8	954	3	US-09-418-641-3	Sequence 3, Appli
22	170	34.5	344	2	US-08-690-011A-28	Sequence 24, Appl
23	169.5	34.4	347	4	US-09-299-495F-24	Sequence 28, Appl
24	169	34.3	462	3	US-08-863-813A-35	Sequence 35, Appl
25	169	34.3	462	3	US-08-676-318A-35	Sequence 35, Appl
26	168	34.1	345	2	US-08-690-011A-20	Sequence 20, Appl
27	167	33.9	344	2	US-09-299-495F-26	Sequence 26, Appl
28	165.5	33.6	347	2	US-08-690-011A-28	Sequence 28, Appl
29	164.5	33.4	341	4	US-09-299-495F-24	Sequence 24, Appl
30	159	32.3	180	2	US-08-256-790-9	Sequence 9, Appli
31	159	32.3	180	2	US-08-256-790-13	Sequence 13, Appl
32	158.5	32.2	729	4	US-09-059-625-47	Sequence 47, Appl
33	157.5	31.9	729	4	US-09-059-625-33	Sequence 33, Appl
34	157.5	31.9	729	4	US-09-059-625-44	Sequence 44, Appl
35	157	31.8	151	2	US-07-801-798-4	Sequence 4, Appli
36	157	31.8	151	5	PCT-US92-10140-4	Sequence 4, Appli
37	156.5	31.7	267	4	US-09-059-625-5	Sequence 5, Appli
38	155.5	31.5	341	2	US-08-690-011A-26	Sequence 26, Appl
39	156.5	25.7	324	4	US-09-299-495F-22	Sequence 22, Appl
40	124.5	25.3	324	2	US-08-690-011A-22	Sequence 22, Appl
41	119	24.1	3489	2	US-08-728-323A-1	Sequence 1, Appli
42	119	24.1	3489	4	US-09-298-568-1	Sequence 1, Appli
43	119	24.1	3489	4	US-09-410-399-1	Sequence 1, Appli
C 44	119	24.1	32207	2	US-08-770-379-20	Sequence 20, Appl
C 45	119	24.1	32207	3	US-08-757-669A-20	Sequence 20, Appl

ALIGNMENTS

RESULT 1

US-08-690-011A-18
; Sequence 18, Application US/08690011A
; Patent No. 5942433

GENERAL INFORMATION:

APPLICANT: VINSON, Charles R.

APPLICANT: KRYLOV, Dmitry

TITLE OF INVENTION: EXTENSION OF A PROTEIN-PROTEIN

TITLE OF INVENTION: INTERACTION SURFACE TO INACTIVATE THE FUNCTION OF A

TITLE OF INVENTION: CELLULAR PROTEIN

NUMBER OF SEQUENCES: 60

CORRESPONDENCE ADDRESS:

ADDRESSEE: MORGAN & FINNEGAN, L.L.P.

STREET: 345 Park Avenue

CITY: New York

STATE: NY

COUNTRY: USA

ZIP: 10154-0053

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.308

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/690, 011A

FILING DATE: 31-JUL-1996

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 60/001,654

FILING DATE: 31-JUL-1995

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 60/018,496

FILING DATE: 29-MAY-1996

ATTORNEY/AGENT INFORMATION:

NAME: Serunian, Leslie A.

REGISTRATION NUMBER: 35,353

REFERENCE/DOCKET NUMBER: 2026-4199US1

APPLICATION NUMBER: US/09/299,495F
FILING DATE: 26-Apr-1999
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/001,654
FILING DATE: 31-JUL-1995
APPLICATION NUMBER: 60/018,496
FILING DATE: 29-MAY-1996
APPLICATION NUMBER: 08/690,011
FILING DATE: 31-JULY-1996
ATTORNEY/AGENT INFORMATION:
NAME: Serunian, Leslie A.
REGISTRATION NUMBER: 35,353
REFERENCE/DOCKET NUMBER: 2026-4199US2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212)758-4800
TELEFAX: (212)751-6849
INFORMATION FOR SEQ ID NO: 16:
SEQUENCE CHARACTERISTICS:
LENGTH: 281 base pairs
TYPE: nucleic acid
STRANDEDNESS: unknown
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
SEQUENCE DESCRIPTION: SEQ ID NO: 16:
US-09-299-495F-16

Alignment Scores:
Pred. No.: 4,28e-45 Length: 281
Score: 441.00 Matches: 88
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 89.45% Indels: 0
DB: 4 Gaps: 0

US-10-059-720-19 (1-97) x US-09-299-495F-16 (1-281)

QY 10 HisMetAlaSerMetThrGlyGlnGlnMetGlyArgAspProAspLeuGlnArg 29
DB 6 CATATGCTAGCATGATGCTGTGGACACAAATGGTGGGATCCTGACCTGGAAACAGCT 65
QY 30 AlaGluLeuAlaArgGluAsnGluLeuGluLysGluAlaGluLeuGln 49
DB 66 GCTGAGAACTGGCCGCGTGAACGAGAGCTGGAAAAAGAGCCGAGAGCTGGAGCAG 125
QY 50 GluAsnAlaGluLeuAlaGluThrAspGlnLeuGluLysGluLysSerAlaLeuGln 69
DB 126 GAAACGCTGAACCTCGAGCGGAGACCACTAGAGATGAGAGTCTGCTTTGCAG 185
QY 70 ThrGluLeuAlaAsnLeuLysGluLysGluLysLeuGluPheIleLeuAlaAlaHis 89
DB 186 ACCGAGATTGCCAACCTGCTGAAGAGAGAGAAAACTAGAGTTTCATCTGCGAGCTCAC 245
QY 90 ArgProAlaCysLysLysProAsp 97
DB 246 CGACCTGCTGCAAGATCCCTGAT 269

RESULT 4

US-08-690-011A-16
Sequence 16, Application US/08690011A
Patent No. 5942433
GENERAL INFORMATION:
APPLICANT: VINSON, Charles R.
APPLICANT: KRYLOV, Dmitry
TITLE OF INVENTION: EXTENSION OF A PROTEIN-PROTEIN
TITLE OF INVENTION: INTERACTION SURFACE TO INACTIVATE THE FUNCTION OF A
TITLE OF INVENTION: CELLULAR PROTEIN
NUMBER OF SEQUENCES: 60
CORRESPONDENCE ADDRESS:
ADDRESSEE: MORGAN & FINNEGAN, L.L.P.
STREET: 345 Park Avenue
CITY: New York
STATE: NY
COUNTRY: USA

ZIP: 10154-0053
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.308
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/690,011A
FILING DATE: 31-JUL-1996

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/001,654
FILING DATE: 31-JUL-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/018,496
FILING DATE: 29-MAY-1996
ATTORNEY/AGENT INFORMATION:
NAME: Serunian, Leslie A.
REGISTRATION NUMBER: 35,353
REFERENCE/DOCKET NUMBER: 2026-4199US1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212)758-4800
TELEFAX: (212)751-6849
INFORMATION FOR SEQ ID NO: 16:
SEQUENCE CHARACTERISTICS:
LENGTH: 281 base pairs
TYPE: nucleic acid
STRANDEDNESS: unknown
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-690-011A-16

Alignment Scores:
Pred. No.: 1,71e-44 Length: 281
Score: 436.00 Matches: 87
Percent Similarity: 98.86% Conservative: 0
Best Local Similarity: 98.86% Mismatches: 1
Query Match: 88.44% Indels: 0
DB: 2 Gaps: 0

US-10-059-720-19 (1-97) x US-08-690-011A-16 (1-281)

QY 10 HisMetAlaSerMetThrGlyGlnGlnMetGlyArgAspProAspLeuGlnArg 29
DB 6 CATATGCTAGCATGATGCTGTGGACACAAATGGTGGGATCCTGACCTGGAAACAGCT 65
QY 30 AlaGluLeuAlaArgGluAsnGluLeuGluLysGluAlaGluLeuGln 49
DB 66 GCTGAGAACTGGCCGCGTGAACGAGAGCTGGAAAAAGAGCCGAGAGCTGGAGCAG 125
QY 50 GluAsnAlaGluLeuAlaGluThrAspGlnLeuGluLysGluLysSerAlaLeuGln 69
DB 126 GAAACGCTGAACCTCGAGCGGAGACCACTAGAGATGAGAGTCTCCTTTGCAG 185
QY 70 ThrGluLeuAlaAsnLeuLysGluLysGluLysLeuGluPheIleLeuAlaAlaHis 89
DB 186 ACCGAGATTGCCAACCTGCTGAAGAGAGAGAAAACTAGAGTTTCATCTGCGAGCTCAC 245
QY 90 ArgProAlaCysLysLysProAsp 97
DB 246 CGACCTGCTGCAAGATCCCTGAT 269

RESULT 5

US-09-299-495F-14
Sequence 14, Application US/09299495F
Patent No. 6361968
GENERAL INFORMATION:
APPLICANT: VINSON, Charles R.
APPLICANT: KRYLOV, Dmitry
TITLE OF INVENTION: EXTENSION OF A PROTEIN-PROTEIN
TITLE OF INVENTION: INTERACTION SURFACE TO INACTIVATE THE FUNCTION OF A
CELLULAR PROTEIN
NUMBER OF SEQUENCES: 64
CORRESPONDENCE ADDRESS:

ADDRESSEE: MORGAN & FINNEGAN, L.L.P.
STREET: 345 Park Avenue
CITY: New York
STATE: NY
COUNTRY: USA
ZIP: 10154-0053
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: MS WORD 97
SOFTWARE: ASCII
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/299,495F
FILING DATE: 26-Apr-1999
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/001,654
FILING DATE: 31-JUL-1995
APPLICATION NUMBER: 60/018,496
FILING DATE: 29-MAY-1996
APPLICATION NUMBER: 09/690,011
FILING DATE: 31-JULY-1996
ATTORNEY/AGENT INFORMATION:
NAME: Serunian, Leslie A.
REGISTRATION NUMBER: 35,353
REFERENCE/DOCKET NUMBER: 2026-4199US2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212)758-4800
TELEFAX: (212)751-6849
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE DESCRIPTION: SEQ ID NO: 14:
US-09-299-495F-14

Alignment Scores:
Pred. No.: 4.6e-34 Length: 324
Score: 350.00 Matches: 72
Percent Similarity: 74.53% Conservative: 7
Best Local Similarity: 67.92% Mismatches: 17
Query Match: 70.99% Indels: 10
DB: 4 Gaps: 2

US-10-059-720-19 (1-97) x US-09-299-495F-14 (1-324)

QY 1 MetAspTyrLysAspAspAspLysHisMetAlaSerMetThrGlyGlnGlnMet 20
DB 3 ATGGACTACAGGACGACGATGACAGCATATGGCTACATGCTGGTGGACAGCAATG 62
QY 21 GlyArgAspProAspLeuGluGlnArgAlaGluGluLeuAlaArgGluAsnGluLeu 40
DB 63 GGTGGGATCCCAAGTGGAAACAGTTATCTCCAGAA---GAAGAAGAGAAAGGAGATC 119
QY 41 GluLysGluAlaGluGluGluGluGlnGluAsnAlaGlu----- 53
DB 120 CGAAGGGAAGATAAGATGGCTGCACGCCAAATGCCCAACCCGAGAGGAGCTGACT 179
QY 54 -----LeuGluAlaGluThrAspGlnLeuGluAspGluLysSerAlaLeuGlnThrGlu 71
DB 180 GATACACTCCAAAGCGAGACAGCACTAGAGTGTAGAGTCTGCTTTCCAGCCGAG 239
QY 72 IleAlaAsnLeuLeuLysGluLysGluLysLeuGluPheIleLeuAlaAlaHisArgPro 91
DB 240 ATTGCCACCTGCTCAGGAGAGGAGAAACTAGATTCTATCTCTGGCAGCTCCCGACT 299
QY 92 AlaCysLysIleProAsp 97
DB 300 GCCTCAAGATCCCTGAT 317

RESULT 6

US-08-690-011A-14
Sequence 14, Application US/08690011A
Patent No. 5942433
GENERAL INFORMATION:
APPLICANT: VINSON, Charles R.
APPLICANT: KRYLOV, Dmitry
TITLE OF INVENTION: EXTENSION OF A PROTEIN-PROTEIN
TITLE OF INVENTION: INTERACTION SURFACE TO INACTIVATE THE FUNCTION OF A
NUMBER OF INVENTION: CELLULAR PROTEIN
NUMBER OF SEQUENCES: 60
CORRESPONDENCE ADDRESS:
ADDRESSEE: MORGAN & FINNEGAN, L.L.P.
STREET: 345 Park Avenue
CITY: New York
STATE: NY
COUNTRY: USA
ZIP: 10154-0053
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30B
CURRENT APPLICATION DATA: US/08/690,011A
APPLICATION NUMBER: US/08/690,011A
FILING DATE: 31-JUL-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/001,654
FILING DATE: 31-JUL-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/018,496
FILING DATE: 29-MAY-1996
ATTORNEY/AGENT INFORMATION:
NAME: Serunian, Leslie A.
REGISTRATION NUMBER: 35,353
REFERENCE/DOCKET NUMBER: 2026-4199US1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212)758-4800
TELEFAX: (212)751-6849
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 318 base pairs
TYPE: nucleic acid
STRANDEDNESS: unknown
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-690-011A-14

Alignment Scores:
Pred. No.: 5.94e-34 Length: 318
Score: 349.00 Matches: 72
Percent Similarity: 74.53% Conservative: 7
Best Local Similarity: 67.92% Mismatches: 17
Query Match: 70.79% Indels: 10
DB: 2 Gaps: 2

US-10-059-720-19 (1-97) x US-08-690-011A-14 (1-318)

QY 1 MetAspTyrLysAspAspAspLysHisMetAlaSerMetThrGlyGlnGlnMet 20
DB 1 ATCGACTACAGGACGACGATGACAGCATATGGCTACATGCTGGTGGACAGCAATG 60
QY 21 GlyArgAspProAspLeuGluGlnArgAlaGluGluLeuAlaArgGluAsnGluLeu 40
DB 61 GGTGGGATCCCAAGTGGAAACAGTTATCTCCAGAA---GAAGAAGAGAAAGGAGATC 117
QY 41 GluLysGluAlaGluGluGluGluGlnGluAsnAlaGlu----- 53
DB 118 CGAAGGGAAGATAAGATGGCTGCAGGCCAAATGCCCAACCCGAGAGGAGCTGACT 177
QY 54 -----LeuGluAlaGluThrAspGlnLeuGluAspGluLysSerAlaLeuGlnThrGlu 71
DB 178 GATACACTCCAGCGGAGACAGCACTAGAGTGTAGAGTCTGCTTTCCAGCCGAG 237

QY 21 GlyArgProAspLeuGluGlnArgAlaGluGluLeuAlaArgGluAsnGluLeu 40
Db 63 GGTGGGATCTGCTGCTGGAACACCTGCTGAGGAACCTGCGCCGTGAAACGAGAGCTG 122
QY 41 GluTyrGluAlaGluGluGluGluGluGluGluGluGluGluGluGluGluGluGlu 60
Db 123 GAAAG 182
QY 61 LeuGluAspGluTyrSerAlaLeuGlnThrGluLeuAlaAsnLeu 75
Db 183 CTGAAACCAACAAACAAACATTCATTCGAGGAGCTAAAGCACTT 227

RESULT 9

US-08-690-011A-32
; Sequence 32, Application US/08690011A
; Patent No. 5942433
; GENERAL INFORMATION:
; APPLICANT: VINSON, Charles R.
; TITLE OF INVENTION: EXTENSION OF A PROTEIN-PROTEIN
; TITLE OF INVENTION: INTERACTION SURFACE TO INACTIVATE THE FUNCTION OF A
; TITLE OF INVENTION: CELLULAR PROTEIN
; NUMBER OF SEQUENCES: 60
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORGAN & FINNEGAN, L.L.P.
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10154-0053
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30B
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/690,011A
; FILING DATE: 31-JUL-1996

; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/001,654
; FILING DATE: 31-JUL-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/018,496
; FILING DATE: 29-MAY-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Serunian, Leslie A.
; REGISTRATION NUMBER: 35,353
; REFERENCE/DOCKET NUMBER: 2026-4199US1
; TELEPHONE: (212)751-6849
; TELEFAX: (212)751-6849
; INFORMATION FOR SEQ ID NO: 32:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 296 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: unknown
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-690-011A-32

Alignment Scores:
Pred. No.: 1,26E-20 Length: 296
Score: 238.00 Matches: 53
Percent Similarity: 75.32% Conservative: 5
Best Local Similarity: 68.83% Mismatches: 17
Query Match: 48.28% Indels: 2
DB: 2 Gaps: 1

US-10-059-720-19 (1-97) x US-08-690-011A-32 (1-296)

QY 11 MetalSerMetThrGlyGlnGlnMetGlyArgAspPro-----AspLeuGluGln 28
Db 1 ATGGCTAGCATGACTGCTGGAGCAGCAAAATGGGTGCGGATCCTGGCGGTGGCTGGAACAA 60

QY 29 ArgAlaGluGluLeuAlaArgGluAsnGluGluLeuGluGluGluGluGluGlu 48
Db 61 CGTCTGAGGAACCTGGCCCGTGAACACGAAAGAGCTGGAACGAGAGAGAGAGAGAGAG 120
QY 49 GlnGluAsnAlaGluLeuGluAlaGluThrAspGlnLeuGluAspGluTyrSerAlaLeu 68
Db 121 CAGAAACACCTGAACTCGAGCAGAGAGTGTGGAGTTGGAAAGTCTGTAATGACCCCTG 180
QY 69 GlnThrGluLeuAlaAsnLeuLeuGluGluGluGluGluGluGluGluGlu 85
Db 181 CGCAAGGAAGTGGAGACAGCTGGAGCGTGAACCTGGACACGCTGGCGGGGTATC 231

RESULT 10

US-09-299-495F-32
; Sequence 32, Application US/09299495F
; Patent No. 6361968
; GENERAL INFORMATION:
; APPLICANT: VINSON, Charles R.
; TITLE OF INVENTION: EXTENSION OF A PROTEIN-PROTEIN
; TITLE OF INVENTION: INTERACTION SURFACE TO INACTIVATE THE FUNCTION OF A
; TITLE OF INVENTION: CELLULAR PROTEIN
; NUMBER OF SEQUENCES: 64
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORGAN & FINNEGAN, L.L.P.
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10154-0053
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: MS WORD 97
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/299,495F
; FILING DATE: 26-Apr-1999

; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/001,654
; FILING DATE: 31-JUL-1995
; APPLICATION NUMBER: 60/018,496
; FILING DATE: 29-MAY-1996
; APPLICATION NUMBER: 08/690,011
; FILING DATE: 31-JULY-1996

; ATTORNEY/AGENT INFORMATION:
; NAME: Serunian, Leslie A.
; REGISTRATION NUMBER: 35,353
; REFERENCE/DOCKET NUMBER: 2026-4199US2

; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212)751-6849
; TELEFAX: (212)751-6849

; INFORMATION FOR SEQ ID NO: 32:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 296 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: unknown
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)

; SEQUENCE DESCRIPTION: SEQ ID NO: 32:
US-09-299-495F-32

Alignment Scores:
Pred. No.: 1,26E-20 Length: 296
Score: 238.00 Matches: 53
Percent Similarity: 75.32% Conservative: 5
Best Local Similarity: 68.83% Mismatches: 17
Query Match: 48.28% Indels: 2
DB: 4 Gaps: 1

US-10-059-720-19 (1-97) x US-09-299-495F-32 (1-296)

Score: 229.00 Matches: 50
Percent Similarity: 67.86% Conservative: 7
Best Local Similarity: 59.52% Mismatches: 17
Query Match: 46.45% Indels: 10
DB: 4 Gaps: 2

US-10-059-720-19 (1-97) x US-09-299-495F-12 (1-262)

QY 23 AspProaspLeuGluInArgAlaGluGluLeuAlaArgGluLeuGluLys 42
| | | | | : : : : :
Db 2 GATCCAGGTGGACAGTATATCCAGAA---GAGAGAGAGAAAGGAGATCCGAGG 58
| | | | | : : : : :
QY 43 GluAlaGluGluLeuGluGluGluAsnAlaGlu----- 53
| | | | | : : : : :
Db 59 GAAAGGAATAGATGGCTGCGCCCAATGCGCAACCGGAGGAGGAGTGTACTGATACA 118
| | | | | : : : : :
QY 54 LeuGluAlaGluThrAspGluLeuGluAspGluLysSerAlaLeuGluInThrGluLeuAla 73
| | | | | : : : : :
Db 119 CTCGAGCGGAGACAGACCAACTAGAGATGAGAAGTCTGTTTGCAGACCGAGATTGCC 178
| | | | | : : : : :
QY 74 AsnLeuLysGluLysGluLysGluPheLeuAlaAlaHisArgProAlaCys 93
| | | | | : : : : :
Db 179 AACCTGCTGAAGGAGAGGAAATACTAGATTCTCTGGCAGCTCACCAGCTGCTGC 238
| | | | | : : : : :
QY 94 LysileProasp 97
| | | | | : : : : :
Db 239 AAGATCCCTGAT 250
| | | | | : : : : :

RESULT 13

US-09-765-298A-1
; Sequence 1, Application US/09765298A
; Patent No. 6582927

; GENERAL INFORMATION:
; APPLICANT: ARONHEIM, AMI

; TITLE OF INVENTION: METHOD FOR DETECTION PROTEIN-PROTEIN INTERACTIONS AND A KIT THERE

; FILE REFERENCE: 108387.01

; CURRENT APPLICATION NUMBER: US/09/765,298A

; CURRENT FILING DATE: 2001-01-22

; PRIOR APPLICATION NUMBER: IL 125456

; PRIOR FILING DATE: 1998-07-22

; PRIOR APPLICATION NUMBER: IL 128017

; PRIOR FILING DATE: 1999-01-12

; NUMBER OF SEQ ID NOS: 31

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 1

; LENGTH: 435

; TYPE: DNA

; ORGANISM: Homo sapiens

US-09-765-298A-1

Alignment Scores:

Pred. No.: 1.43e-16 Length: 435
Score: 206.00 Matches: 41
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 41.78% Indels: 0
DB: 4 Gaps: 0

US-10-059-720-19 (1-97) x US-09-765-298A-1 (1-435)

QY 57 GluThrAspGluLeuGluAspGluLysSerAlaLeuGluInThrGluLeuAlaAsnLeuLeu 76
| | | | | : : : : :
Db 220 GAGACAGCACTAGAGATGAGAAGTCTGTTTGCAGACCGAGATTGCCAAGTCTGTG 279
| | | | | : : : : :
QY 77 LysGluLysGluLysGluPheLeuAlaAlaHisArgProAlaCysLysilePro 96
| | | | | : : : : :
Db 280 AAGGAGAGGAAATACTAGATTCTCTGGCAGCTCACCAGCTGCTGCAGATCCCT 339
| | | | | : : : : :

QY 97 Asp 97

| | | | |

Db 340 GAT 342

| | | | |

RESULT 14

US-08-306-691B-44
; Sequence 44, Application US/08306691B
; Patent No. 5734039

; GENERAL INFORMATION:

; APPLICANT: Calabretta, Bruno

; APPLICANT: Skorski, Tomasz

; TITLE OF INVENTION: ANTISENSE

; TITLE OF INVENTION: OLIGONUCLEOTIDES TARGETING COOPERATING ONCOGENES

; NUMBER OF SEQUENCES: 55

; CORRESPONDENCE ADDRESS:

; ADDRESS: Seidel, Gonda, Lavorgha & Monaco, P.C.

; STREET: Two Penn Center, Suite 1800

; CITY: Philadelphia

; STATE: Pennsylvania

; COUNTRY: U.S.A.

; ZIP: 19102

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette, 3.50 inch, 720 KB

; COMPUTER: IBM PS/2

; OPERATING SYSTEM: MS-DOS

; SOFTWARE: WordPerfect 5.1

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/306,691B

; FILING DATE: September 15, 1994

; CLASSIFICATION: 514

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER:

; FILING DATE:

; ATTORNEY/AGENT INFORMATION:

; NAME: Monaco, Daniel A.

; REGISTRATION NUMBER: 30,480

; REFERENCE/DOCKET NUMBER: 8321-8

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (215) 568-8383

; TELEFAX: (215) 568-5549

; TELEX: No. 5734039e

; INFORMATION FOR SEQ ID NO: 44:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 3565 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: double

; TOPOLOGY: linear

US-08-306-691B-44

Alignment Scores:

Pred. No.: 1.83e-15 Length: 3565
Score: 206.00 Matches: 41
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 41.78% Indels: 0
DB: 1 Gaps: 0

US-10-059-720-19 (1-97) x US-08-306-691B-44 (1-3565)

QY 57 GluThrAspGluLeuGluAspGluLysSerAlaLeuGluInThrGluLeuAlaAsnLeuLeu 76
| | | | | : : : : :
Db 2088 GAGACAGCACTAGAGATGAGAAGTCTGTTTGCAGACCGAGATTGCCAAGTCTGTG 2147
| | | | | : : : : :
QY 77 LysGluLysGluLysGluPheLeuAlaAlaHisArgProAlaCysLysilePro 96
| | | | | : : : : :
Db 2148 AAGGAGAGGAAATACTAGATTCTCTGGCAGCTCACCAGCTGCTGCAGATCCCT 2207
| | | | | : : : : :

QY 97 Asp 97

| | | | |

Db 2208 GAT 2210

| | | | |

RESULT 15

PCT-US93-06251-20

; Sequence 20, Application PC/TUS9306251

; GENERAL INFORMATION:

; APPLICANT: Wickstrom, Eric and Rife, Jason P.

; TITLE OF INVENTION: Trivalent Synthesis of Oligonucleotides Containing

; TITLE OF INVENTION: Stereospecific Alkylphosphonates and Arylphosphonates

NUMBER OF SEQUENCES: 93
CORRESPONDENCE ADDRESS:
ADDRESSEE: SCULLY, SCOTT, MURPHY & PRESSER
STREET: 400 Garden City Plaza
CITY: Garden City
STATE: NY
COUNTRY: USA
ZIP: 11530
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US93/06251
FILING DATE: 19930630
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Digiglio, Frank S.
REGISTRATION NUMBER: 31,346
REFERENCE/DOCKET NUMBER: 8586
TELECOMMUNICATION INFORMATION:
TELEPHONE: 516-742-4343
TELEFAX: 516-742-4366
TELEX: 230 901 SANS UR
INFORMATION FOR SEQ ID NO: 20:
SEQUENCE CHARACTERISTICS:
LENGTH: 3565 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
PCT-US93-06251-20

Alignment Scores:
Pred. No.: 1.83e-15 Length: 3565
Score: 206.00 Matches: 41
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 41.78% Indels: 0
DB: 5 Gaps: 0

US-10-059-720-19 (1-97) x PCT-US93-06251-20 (1-3565)

QY	57	GlutThrAspGlnLeuGluAspGluLysSerAlaLeuGlnThrGluIleAlaAsnLeuLeu	76
Db	2088	GGACAGACACCACTAGAGAGTCTCTTTGCGACCGAGATGCCACCTGCTG	2147
QY	77	LysGluLysGluLysLeuGluPheIleLeuAlaAlaHisArgProAlaCysLysIlePro	96
Db	2148	AAGGAGAGGAGAAACTAGAGTTTCCTCGCAGCTCACCAGCCTGCTGCAAGATCCCT	2207
QY	97	Asp	97
Db	2208	GAT	2210

Search completed: November 16, 2003, 02:57:53
Job time : 58 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: November 16, 2003, 00:59:01 ; Search time 230 Seconds
(without alignments)
1138.459 Million cell updates/sec

Title: US-10-059-720-19

Perfect score: 493
Sequence: 1 MDYKDDDDKHWSMTGQQM.....EKEKLEFILAHRRACKIPD 97

Scoring table: BIOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+p2n.model -DEV=xlh
-O=/cgn2.1/USFTO.spool -SUFFIX=ring -MINMATCH=0.1 -LOOPEXT=0
-DB=N Geneseq 19Jun03 -QPMF=fastap -START=1 -END=1 -MATRIX=blomsum62 -TRANS=human40.cdi
-LIST=45 -DOALIGN=p2n -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODS=LOCAL -OUTFMT=p2n -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10059720 -CGEN 1 1 0 -@runat 14112003 184958 17880 -NCPU=6 -ICPU=3
-NO MMAP -LARGQUERY -NEG SCORES=0 -WAIT -DSBBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N Geneseq 19Jun03.*
1: /SIDSI/cgdata/geneseq/geneseq-emb1/NA1980.DAT.*
2: /SIDSI/cgdata/geneseq/geneseq-emb1/NA1981.DAT.*
3: /SIDSI/cgdata/geneseq/geneseq-emb1/NA1982.DAT.*
4: /SIDSI/cgdata/geneseq/geneseq-emb1/NA1983.DAT.*
5: /SIDSI/cgdata/geneseq/geneseq-emb1/NA1984.DAT.*
6: /SIDSI/cgdata/geneseq/geneseq-emb1/NA1985.DAT.*
7: /SIDSI/cgdata/geneseq/geneseq-emb1/NA1986.DAT.*
8: /SIDSI/cgdata/geneseq/geneseq-emb1/NA1987.DAT.*
9: /SIDSI/cgdata/geneseq/geneseq-emb1/NA1988.DAT.*
10: /SIDSI/cgdata/geneseq/geneseq-emb1/NA1989.DAT.*
11: /SIDSI/cgdata/geneseq/geneseq-emb1/NA1990.DAT.*
12: /SIDSI/cgdata/geneseq/geneseq-emb1/NA1991.DAT.*
13: /SIDSI/cgdata/geneseq/geneseq-emb1/NA1992.DAT.*
14: /SIDSI/cgdata/geneseq/geneseq-emb1/NA1993.DAT.*
15: /SIDSI/cgdata/geneseq/geneseq-emb1/NA1994.DAT.*
16: /SIDSI/cgdata/geneseq/geneseq-emb1/NA1995.DAT.*
17: /SIDSI/cgdata/geneseq/geneseq-emb1/NA1996.DAT.*
18: /SIDSI/cgdata/geneseq/geneseq-emb1/NA1997.DAT.*
19: /SIDSI/cgdata/geneseq/geneseq-emb1/NA1998.DAT.*
20: /SIDSI/cgdata/geneseq/geneseq-emb1/NA1999.DAT.*
21: /SIDSI/cgdata/geneseq/geneseq-emb1/NA2000.DAT.*
22: /SIDSI/cgdata/geneseq/geneseq-emb1/NA2001A.DAT.*
23: /SIDSI/cgdata/geneseq/geneseq-emb1/NA2001B.DAT.*
24: /SIDSI/cgdata/geneseq/geneseq-emb1/NA2002.DAT.*
25: /SIDSI/cgdata/geneseq/geneseq-emb1/NA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed,

and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match %	Length	DB	ID	Description
1	493	100.0	294	18	AAT84348	CMV400-4heptadFos
2	436	88.4	281	18	AAT84347	4heptadFos protein
3	349	70.8	318	18	AAT84346	CMV500-FosBZIP(NO)
4	288	58.4	264	18	AAT84344	DNA encoding CMV50
5	274	55.6	418	24	ABK95199	PCR template conta
6	274	55.6	5172	25	ACC44984	Vector CMV-FosCBPz
7	274	55.6	5172	25	ACC45014	Vector CMV-FosCBP
8	245	49.7	403	25	ABX45369	Bovine EST associa
9	241	48.9	1143	24	ABK50445	Human v-fos FBJ mu
10	241	48.9	2084	24	ABV94327	Breast carcinoma r
11	241	48.9	2103	20	AAZ41352	Human normal uteru
12	241	48.9	2180	24	ABQ54467	Human ovarian anti
13	238	48.3	296	18	AAT84355	4Heptad-F zipper D
14	236	47.9	399	25	ABX42690	Bovine EST associa
15	229	46.5	262	18	ABT84345	Human c-Fos protei
16	216	43.8	273	22	AAE79982	Nucleotide sequenc
17	216	43.8	415	25	ABX49741	Bovine EST associa
18	210	42.6	1176	25	ABX62223	Human activated T
19	206	41.8	642	22	ABA49894	Human breast cell
20	206	41.8	642	22	ABA67812	Human foetal liver
21	206	41.8	642	22	ABA34871	Probe #13337 for g
22	206	41.8	642	22	AAK16224	Human brain expres
23	206	41.8	642	22	AAK41368	Human bone marrow
24	206	41.8	642	22	AAI22734	Probe #12667 for g
25	206	41.8	642	22	AAI48034	Probe #16720 used
26	206	41.8	642	22	AAI08402	Probe #8393 used t
27	206	41.8	642	24	ABS15995	Human genome-deriv
28	206	41.8	1946	22	ABA44752	Human breast cell
29	206	41.8	1946	22	ABA55207	Human foetal liver
30	206	41.8	1946	22	ABA24954	Probe #3420 for ge
31	206	41.8	1946	22	AAK03468	Human bone marrow
32	206	41.8	1946	22	AAK28922	Human brain expres
33	206	41.8	1946	22	AAI13509	Probe #3442 for ge
34	206	41.8	1946	22	AAI34869	Probe #3555 used t
35	206	41.8	1946	22	AAI03394	Probe #3385 used t
36	206	41.8	1946	24	ABS03451	Human genome-deriv
37	206	41.8	2227	25	ABX62920	Human activated T
38	206	41.8	3548	15	AAQ63814	c-fos gene. Mus m
39	206	41.8	3585	19	AAV20466	Human c-fos oncoge
40	206	41.8	3585	21	AAE21092	Human low adenosin
41	206	41.8	3585	21	AAA34370	Human adenosine re
42	206	41.8	3585	24	ABK84488	Human cDNA differe
43	206	41.8	3585	24	ABK64754	Human benign prost
44	206	41.8	3585	24	ABL62456	Colon adenocarcino
45	206	41.8	3585	24	ABL68866	Kidney cancer rela

ALIGNMENTS

RESULT 1
AAT84348
ID AAT84348 standard; DNA; 294 BP.
XX AAT84348;
AC AAT84348;
XX
XX 12-NOV-1997 (first entry)
XX
XX CMV400-4heptadFos leucine zipper protein DNA.
XX
XX DNA binding protein; RNA binding protein; amphipathic peptide;
XX acidic extension peptide; Gene control; gene regulation;
XX transcription; dominant negative protein; c-Fos; cancer;
XX drug therapy; drug design; ss.
XX
XX Chimeric Homo sapiens;
OS Chimeric bacteriophage phi-10;

OS Chimeric synthetic.

XX WO9705249-A2.
XX 13-FEB-1997.
XX 31-JUL-1996; 96WO-US12590.
XX 31-JUL-1996; 96US-0001654.
XX 31-JUL-1995; 95US-0001654.
XX 29-MAY-1996; 96US-0018496.
XX (KRYL/) KRYLOV D.
PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
PA (VINS/) VINSON C R.
XX Krylov D, Vinson CR;
XX WPI; 1997-145687/13.
DR P-PSDB; AAW00949.
XX New nucleic acid binding proteins - having an acidic amino acid
FT sequence extension at the amino-terminus, to increase ability to
FT regulate gene transcription, useful e.g. in cancer therapeutics
XX Claim 16; Page 87; 144pp; English.
XX This DNA sequence codes for CMV500-heptadFos leucine zipper
CC (AAW00948), a modified 4-heptad acidic binding protein (NABP) that
CC includes an appended 4-heptad acidic extension. Claimed NABPs such
CC as Fos having acidic peptide extensions are capable of regulating
CC the function of a target nucleic acid or gene to which they are
CC bound, and act as potent dominant-negative regulators of gene
CC transcription, cell growth and cell proliferation. They can be
CC used in cancer therapeutics, to treat diseases caused by eukaryotic
CC microorganisms or by viruses, and as tools for drug development,
CC rational drug design, and drug and gene therapies. They have an
CC extended protein interaction surface or multimerisation or
CC dimerisation interface that increases the stability of complexes
CC formed.

XX Sequence 294 BP; 96 A; 65 C; 88 G; 45 T; 0 other;

Alignment Scores:

Pred. No.: 3,348-48 Length: 294

Score: 493.00 Matches: 97

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 100.00% Indels: 0

DB: 18 Gaps: 0

US-10-059-720-19 (1-97) x AAT84347 (1-294)

QY 1 MetAspTyrLysAspAspAspLysHisMetAlaSerMetThrGlyGlnGlnMet 20
DB 1 ATGGACTACAGAGCAGCAGTACCAAGCATATGCTAGCATGACTGTGACGAAATG 60
QY 21 GlyArgAspProAspLeuGluGlnArgAlaGluGluLeuAlaArgGluAsnGluGluLeu 40
DB 61 GGTGGGATCTCTGACTGGACACAGCTCTGAGGACTGGCCCTGAAACAGAGCTG 120
QY 41 GluLysGluAlaGluGluLeuGluGlnAlaGluLeuGluAlaGluThrAspGln 60
DB 121 GAAAGAGCCGAGAGCTGGAGCAGGAAACGCTGAACTCGAGGCGGAGACAGACCA 180
QY 61 LeuGluAspGluLysSerAlaLeuGlnThrGluLeuAlaAsnLeuLysGluLysGlu 80
DB 181 CTAGAGATGAGAGTCTGCTTTGAGACCGAGATTGCCACCTGCTGAGAGGAGGAA 240
QY 81 LysLeuGluPheIleuAlaAlaHisArgProAlaCysLysIleProAsp 97
DB 241 AAACCTAGATTCTGCGCAGCTCAGCGAGCTGCTGCAAGATCCCTGAT 291

RESULT 2

AAT84347

ID AAT84347 standard; DNA; 281 BP.

XX AAT84347;

AC AAT84347;

XX 11-NOV-1997 (first entry)

DE 4HeptadFos protein DNA.

XX DNA binding protein; RNA binding protein; amphipathic peptide;

XX acidic extension peptide; gene control; gene regulation;

XX transcription; dominant negative protein; c-Fos; cancer;

XX drug therapy; drug design; ss.

XX Chimeric Homo sapiens;

OS Chimeric bacteriophage phi-10;

OS Chimeric synthetic.

XX Key Location/Qualifiers

FT CDS 9...272

FT /*tag= a

XX WO9705249-A2.

XX 13-FEB-1997.

XX 31-JUL-1996; 96WO-US12590.

XX 31-JUL-1996; 96US-0001654.

XX 31-JUL-1995; 95US-0001654.

XX 29-MAY-1996; 96US-0018496.

XX (KRYL/) KRYLOV D.

PA (USSH) US DEPT HEALTH & HUMAN SERVICES.

PA (VINS/) VINSON C R.

XX Krylov D, Vinson CR;

XX WPI; 1997-145687/13.

XX P-PSDB; AAW00948.

XX New nucleic acid binding proteins - having an acidic amino acid

XX sequence extension at the amino-terminus, to increase ability to

XX regulate gene transcription, useful e.g. in cancer therapeutics

XX Claim 16; Page 86; 144pp; English.

XX This DNA sequence codes for 4heptadFos protein (AAW00948), a

XX modified Fos nucleic acid binding protein (NABP) that includes

XX an appended 4-heptad acidic extension. Claimed NABPs such as

XX Fos that have acidic peptide extensions are capable of regulating

XX the function of a target nucleic acid or gene to which they are

XX bound, and act as potent dominant-negative regulators of gene

XX transcription, cell growth and cell proliferation. They can be

XX used in cancer therapeutics, to treat diseases caused by eukaryotic

XX microorganisms or by viruses, and as tools for drug development,

XX rational drug design, and drug and gene therapies. They have an

XX extended protein interaction surface or multimerisation or

XX dimerisation interface that increases the stability of complexes

XX formed.

XX Sequence 281 BP; 90 A; 63 C; 80 G; 48 T; 0 other;

Alignment Scores:

Pred. No.: 1,166-41 Length: 281

Score: 436.00 Matches: 87

Percent Similarity: 98.86% Conservative: 0

Best Local Similarity: 98.86% Mismatches: 1

Query Match: 88.44% Indels: 0

DB: 18 Gaps: 0

US-10-059-720-19 (1-97) x AAT84347 (1-281)

QY 10 HisMetAlaSerMetThrGlyGlyGlnGlnMetGlyArgAspProAspLeuGlnArg 29
 DB 6 CATATGGCTAGCATGACTGTGGACGACAAATGGGTGGGATCTCGACCTGGAAACAGT 65
 QY 30 AlaGluLeuAlaArgGluAengluGluLeuGluLeuGluLeuGluLeuGluLeu 49
 DB 66 GCTGAGAACTGGCCCTGGAAACGAGAGCTGGAAAGAGGCGGAGAGCTGGAGCG 125
 QY 50 GluAsnAlaGluLeuGluAlaGluThrAspGlnLeuGluAspGluLeuSerAlaLeuGln 69
 DB 126 GAAACGCTGAACCTCGAGGGGAGACAGACCAACTAGAGATGAGAAGTCTCTTTGCAG 185
 QY 70 ThrGluLeuAlaSerLeuLeuGluLeuGluLeuGluLeuGluLeuGluLeuAlaHis 89
 DB 186 ACCGAGATGCAACCTGCTGAGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 245
 QY 90 ArgProAlaCysLysIleProAsp 97
 DB 246 CGACCTGCTGCAGATCCCTGAT 269

RESULT 3

AAT84346

ID AAT84346 standard; DNA; 318 BP.

AC AAT84346;

DT 11-NOV-1997 (first entry)

XX CMV500-FosBZIP(MO) DNA.

XX DNA binding protein; RNA binding protein; amphipathic peptide;
 KW acidic extension peptide; gene control; gene regulation;
 KW transcription; dominant negative protein; c-Fos; cancer;
 KW drug therapy; drug design; ss.

XX Chimeric Homo sapiens;
 OS Chimeric bacteriophage phi 10;
 OS Chimeric synthetic.

XX Key Location/Qualifiers
 FH CDS 1..318
 FT /*tag= a
 FT /transl_except= (pos:1..123, aa:Arg)
 FT /transl_except= (pos:121..123, aa:Arg)

XX WO9705249-A2.

XX 13-FEB-1997.

XX 31-JUL-1996; 96WO-US12590.

XX 31-JUL-1996; 96US-0001654.

XX 31-JUL-1996; 96US-0001654.

XX 29-MAY-1996; 96US-0018496.

XX (KRYL/) KRYLOV D.
 PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
 PA (VINS/) VINSON C R.

XX Krylov D, Vinson CR;
 XX WPI; 1997-145687/13.
 DR P-PSDB; AAW00947.

XX New nucleic acid binding proteins - having an acidic amino acid
 PT sequence extension at the amino-terminus, to increase ability to
 PT regulate gene transcription, useful e.g. in cancer therapeutics

XX Claim 16; Page 85; 144pp; English.

XX This DNA sequence codes for CMV500-FosBZIP(MO) (AAW00947), a
 CC modified nucleic acid binding protein (NABP). Claimed NABPs such

CC as Fos that have acidic peptide extensions are capable of regulating
 CC the function of a target nucleic acid or gene to which they are
 CC bound, and act as potent dominant-negative regulators of gene
 CC transcription, cell growth and cell proliferation. They can be
 CC used in cancer therapeutics, to treat diseases caused by eukaryotic
 CC microorganisms or by viruses, and as tools for drug development.
 CC rational drug design, and drug and gene therapies. They have an
 CC extended protein interaction surface or multimerisation or
 CC dimerisation interface that increases the stability of complexes
 CC formed.

XX SQ Sequence 318 BP; 110 A; 69 C; 90 G; 49 T; 0 other;

Alignment Scores:

Pred. No.: 1.4e-31 Length: 318
 Score: 349.00 Matches: 72
 Percent Similarity: 74.53% Conservative: 7
 Best Local Similarity: 67.92% Mismatches: 17
 Query Match: 70.79% Indels: 10
 DB: 18 Gaps: 2

US-10-059-720-19 (1-97) x AAT84346 (1-318)

QY 1 MetAspTyrIysAspAspAspAspLysHisMetAlaSerMetThrGlyGlyGlnMet 20
 DB 1 ATCGACTACACAGGACGACGATGACAAAGCATATGGCTAGCATGCTGCTGGACGAAATG 60
 QY 21 GlyArgAspProAspLeuGlnArgAlaGluLeuAlaArgGluAsnGluLeu 40
 DB 61 GGTGGGTTCCTCCAGGTGGACAGTATCTCCAGAA---GAGAGAGAGAAAGGAGATC 117
 QY 41 GluLysGluAlaGluLeuGluGluGluGluGluGluGluGluGluGluGluGlu 53
 DB 118 CGAAAGGAAAGGAATAAGATGGCTGCGACCAATGCGCAACCGGAGGAGGAGCTGACT 177
 QY 54 -----LeuGluAlaGluThrAspGlnLeuGluAspGluLysSerAlaLeuGlu 71
 DB 178 GATACACTCCGCGGAGACAGACCACTAGAGATGAGAGTCTCTTTGACGACCGAG 237
 QY 72 IleAlaSerLeuLysGluLysGluLysGluLysLeuGluPheIleLeuAlaHisArgPro 91
 DB 238 ATTGCCAACCTGCTGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 297
 QY 92 AlaCysLysIleProAsp 97
 DB 298 GCCTGCAGATCCCTGAT 315

RESULT 4

AAT84344

ID AAT84344 standard; DNA; 264 BP.

AC AAT84344;

XX 11-NOV-1997 (first entry)

XX DNA encoding CMV500-4heptadCREB (New4hepCREB).
 XX DNA binding protein; RNA binding protein; amphipathic peptide;
 KW acidic extension peptide; gene control; gene regulation;
 KW transcription; dominant negative protein; CREB; cancer;
 KW drug therapy; drug design; CMV; ss.

XX Chimeric Homo sapiens;
 OS Chimeric cytomegalovirus;
 OS Chimeric synthetic.

XX Key Location/Qualifiers
 FH CDS 1..255
 FT /*tag= a
 FT /transl_except= (pos:229..231, aa:Glu)

XX WO9705249-A2.

```

PD 13-FEB-1997.
XX
PF 31-JUL-1996; 96WO-US12590.
XX
PR 31-JUL-1996; 96US-0001654.
XX
PR 31-JUL-1996; 96US-0001654.
XX
PR 29-MAY-1996; 96US-0018496.
XX
XX (KRYL/) KRYLOV D.
PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
PA (VINS/) VINSON C R.
XX
PI Krylov D, Vinson CR;
XX
XX WPI; 1997-145687/13.
XX
DR P-PSDB; RAW00945.
XX
XX New nucleic acid binding proteins - having an acidic amino acid
PT sequence extension at the amino-terminus, to increase ability to
PT regulate gene transcription, useful e.g. in cancer therapeutics
XX
XX Claim 16; Page 83; 144pp; English.
XX
XX This DNA sequence codes for CMV500-4heptadCREB (AAW00945), or
CC NewhepCREB, which has a 4heptad appended acidic extension.
CC Claimed nucleic acid binding proteins (NABPs) such as CREB that
CC have acidic peptide extensions are capable of regulating the
CC function of a target nucleic acid or gene to which they are bound,
CC and act as potent dominant-negative regulators of gene
CC transcription, cell growth and cell proliferation. They can be
CC used in cancer therapeutics, to treat diseases caused by eukaryotic
CC microorganisms or by viruses, and as tools for drug development,
CC rational drug design, and drug and gene therapies. They have an
CC extended protein interaction surface or multimerisation or
CC dimerisation interface that increases the stability of complexes
CC formed.
XX
XX Sequence 264 BP; 93 A; 52 C; 76 G; 43 T; 0 other;
SQ
Alignment Scores:
Pred. No.: 1.19e-24 Length: 264
Score: 288.00 Matches: 59
Percent Similarity: 82.67% Conservative: 3
Best Local Similarity: 78.67% Mismatches: 13
Query Match: 58.42% Indels: 0
DB: 18 Gaps: 0
US-10-059-720-19 (1-97) x AAT84344 (1-264)
QY 1 MetAspTyrLysAspAspAspLysHisMetAlaSerMetThrGlyGlnGlnMet 20
DB 1 ATGGACTACAGGACGACGATGACAGCATATGCTAGCATGCTGGTGCACAGCAATG 60
QY 21 GlyArgAspProAspLeuGluGluArgAlaGluGluLeuAlaGluGluLeu 40
DB 61 GTCGGGATCCGGACCTCGAACAACGCTGAGGAACTGGCCGCGTGAACACGAAAGAGCTG 120
QY 41 GluLysGluAlaGluGluGluGluGluGluGluGluGluGluGluGluGluGlu 60
DB 121 GAAAAGAGGCGGAGAGCTGAGCAGCAACTGCGCAGCACTCGAGAACAGAGTGGCAGTG 180
QY 61 LeuGluAspGluLysSerAlaLeuGlnThrGluLeuAlaLeu 75
DB 181 CTTGAAACCAACAAACAAACATTGATTGAGGAGCTAAAGACATT 225
RESULT 5
ID ABK95199
XX ABK95199 standard; DNA; 418 BP.
XX
AC ABK95199;
XX
DT 24-SEP-2002 (first entry)
XX

```

PCR template containing part of c-fos: c-fos (pSPAM).

C-terminal modified protein; protein interaction detection;
proteome analysis; protein-nucleic acid interaction; ss.
Synthetic.

WO200246395-A1.
13-JUN-2002.
07-DEC-2001; 2001WO-JP10731.
07-DEC-2000; 2000JP-0373105.
(UYKE-) UNIV KEIO.
Yanagawa H, Doi N, Miyamoto E, Takashima H, Oyama R;
WPI; 2002-500446/53.
Production of C-terminal modified proteins with nucleotide-linker
containing modifying agents and translation templates, useful for
detecting protein interaction in functional analysis of genes e.g. in
genome projects
Example 2; Page 84; 95pp; Japanese.
The invention relates to an agent for modifying the C-terminal of a
protein comprising an acceptor region with a group capable of binding to
a protein through a transpeptidation reaction in a protein translation
system, and a modifying region containing a non-radioactive modifier
linked to a part of the acceptor region via a nucleotide linker. The
modified proteins are useful for detecting protein interaction in
functional analysis of genes e.g. in genome projects, as well as protein-
nucleic acid interaction in large quantities in high-throughput screening
when studying biological molecules such as proteins and nucleic acids in
genome function or proteome analysis. The modified proteins can be
conveniently and quickly applied in studying protein interactions, with
improved efficiency. ABK95189-ABK95225 represent PCR primers templates
used in examples of the invention.

Sequence 418 BP; 127 A; 91 C; 115 G; 85 T; 0 other;
SQ

Alignment Scores:
Pred. No.: 8.39e-23 Length: 418
Score: 274.00 Matches: 64
Percent Similarity: 63.89% Conservative: 5
Best Local Similarity: 59.26% Mismatches: 15
Query Match: 55.58% Indels: 24
DB: 24 Gaps: 3

US-10-059-720-19 (1-97) x ABK95199 (1-418)

QY 11 MetAlaSerMetThrGlyGlnGlnMetGlyArgAspProAspLeuGluGlnArgAla 30
DB 91 ATGGCTAGCATGCTGCTGCACAGCAATGCTGGTGGCGA-----TCGGCGAGCGG 141
QY 31 GluGluLeuAlaArgGluAsn-----GluGluLeuGluGluGluAlaGluGluGlu 48
DB 142 CAGAGCATCGCGAGGGGCGCAAGATAGAGACGCTATCTCTCGAAGAGAGAGAACCG 201
QY 49 GlnGluAsnAlaGlu-----GlnGluGluGluGluGluGluGluGluGluGlu 53
DB 202 AGAATCCGAGGAGACGGAATAGATGCTGCGACCCCACTGCCGATCGGAGGAGGAG 261
QY 54 -----LeuGluAlaGluThrAspGlnLeuGluAspGluLysSerAlaLeuGln 69
DB 262 CTGACAGATACACTCCACGCGGAGACAGATCAACTTGAAGATGAGAGTCTCGGTGCG 321
QY 70 ThrGluLeuAlaAsnLeuLeuLysGluLysGluLysGluLysGluLysGluLysGlu 89
DB 322 ACTGAGATTGCCAATCTGCTGAAGAGAGAGGAGAAACTGGAATTTATTTTGGCAGCCAC 381

QY 90 ArgProAlaCysLysIleProAsp 97
 DB 382 CGACCTGCTGCAAGATCCCGAT 405

RESULT 6
 ACC44984
 ID ACC44984 standard; DNA; 5172 BP.
 AC ACC44984;
 XX
 XX
 XX 09-JUN-2003 (first entry)
 DT
 DE Vector CMV-FosCBPz2 nucleotide sequence SEQ ID NO:1.
 KW Interaction; detection; cell-free translation system; protein function;
 KW gene function; medicine; food science; environmental monitoring;
 KW analysis; gene; ds.
 XX
 XX Synthetic.
 XX
 XX W02003014734-A1.
 PN
 XX
 XX 20-FEB-2003.
 PD
 XX
 XX 07-AUG-2002; 2002WO-JP08078.
 PF
 XX
 XX 07-AUG-2001; 2001JP-0239175.
 PR
 XX
 XX 12-JUN-2002; 2002JP-0172129.
 PR
 XX
 XX (UYKE-) UNIV KEIO.
 PA
 XX
 XX Yanagawa H, Miyamoto E, Matsumura N, Doi N, Tateyama S;
 PI Ishizaka M, Hozisawa K;
 XX
 XX WPI; 2003-248336/24.
 DR
 XX Bait-prey method for detecting interaction between substances and
 PT proteins in cell-free translation system in vitro, applicable in
 PT screening interacting substances, and analysis of protein and gene
 PT functions -
 XX
 XX Example 1; Page 86-88; 114pp; Japanese.
 PS
 XX The present invention describes a method for detecting the interaction
 CC between a bait and a prey, which comprises bringing them into contact
 CC then detecting the complex which is formed, in which a fused protein of
 CC the prey with a protein applicable as a label for detection is formed
 CC while the bait is modified for separation, and mRNA that encodes the
 CC fused protein containing the prey is translated in a cell-free
 CC translation system in the presence of the bait for interaction with the
 CC produced prey before separating and detecting via the label. The method
 CC is useful for detecting interaction between substances and proteins in a
 CC cell-free translation system in vitro, which is applicable in screening
 CC interacting substances, and analysis of interactions between proteins
 CC and protein-nucleic acids as well as protein and gene functions, useful
 CC in medicine, food science and environmental monitoring. The method can be
 CC performed for global detection layerwise, even of weak interactions at
 CC high sensitivity, in the presence of a bait or prey, or full-length bait,
 CC in a cell-free translation system. But, there is no need for large-scale
 CC expression and purification of such bait by other means after production
 CC in Escherichia coli to greatly shorten the operating time. The present
 CC sequence represents a vector nucleotide sequence which is used in an
 CC example from the present invention.
 XX
 XX Sequence 5172 BP; 1365 A; 1307 C; 1291 G; 1209 T; 0 other;
 SQ

DB: 25 Gaps: 3

US-10-059-720-19 (1-97) x ACC44984 (1-5172)

QY 11 MetAlaSerMetThrGlyGlnGlnMetGlyArgAspProAspLeuGluGlnArgAla 30
 DB 794 ATGGCTAGCATGACTGGTGGACAGCAATGGGTGCGGA-----TCCGGCAGACGG 844
 QY 31 GluGluLeuAlaArgGluAsn-----GluGluLeuGluLysGluAlaGluGluLeuGlu 48
 DB 845 CAGAGCATCGGCAAGAGGGGCAAGTAGCAGCTATCTCTGAGAGAGAGAGAGAGACGG 904
 QY 49 GlnGluAsnAlaGlu-----53
 DB 905 AGAATCCGAGGAGGACCGGATAGATGCTGCGAGCCAGTCCCGAATCGGAGGAGCGAG 964
 QY 54 -----LeuGluAlaGluThrAspGlnLeuGluAspGluLysSerAlaLeuGln 69
 DB 965 CTGACAGATACACTCCAGCGGAGACAGATCACTTGAAGATGAGAACTCTGCGTTGCAG 1024
 QY 70 ThrGluIleAlaAsnLeuLysGluLysGluLysLeuGluPheIleLeuAlaAlaHis 89
 DB 1025 ACTGAGATTGCCAATCTGCTGAAAGAGAGAGAGAACTGGAGTTTATTTCGAGCCAC 1084
 QY 90 ArgProAlaCysLysIleProAsp 97
 DB 1085 CGACCTGCTGCAAGATCCCGAT 1108

RESULT 7
 ACC45014
 ID ACC45014 standard; DNA; 5172 BP.
 XX
 XX ACC45014;
 XX
 XX 09-JUN-2003 (first entry)
 DT
 XX
 XX Vector pCMV-FosCBPz2 nucleotide sequence SEQ ID NO:31.
 XX
 XX Interaction; detection; cell-free translation system; protein function;
 KW gene function; medicine; food science; environmental monitoring;
 KW analysis; gene; ds.
 XX
 XX Synthetic.
 XX
 XX W02003014734-A1.
 PN
 XX
 XX 20-FEB-2003.
 PD
 XX
 XX 07-AUG-2002; 2002WO-JP08078.
 PF
 XX
 XX 07-AUG-2001; 2001JP-0239175.
 PR
 XX
 XX 12-JUN-2002; 2002JP-0172129.
 PR
 XX
 XX (UYKE-) UNIV KEIO.
 PA
 XX
 XX Yanagawa H, Miyamoto E, Matsumura N, Doi N, Tateyama S;
 PI Ishizaka M, Hozisawa K;
 XX
 XX WPI; 2003-248336/24.
 DR
 XX Bait-prey method for detecting interaction between substances and
 PT proteins in cell-free translation system in vitro, applicable in
 PT screening interacting substances, and analysis of protein and gene
 PT functions -
 XX
 XX Example 5; Page 103-105; 114pp; Japanese.
 PS
 XX The present invention describes a method for detecting the interaction
 CC between a bait and a prey, which comprises bringing them into contact
 CC then detecting the complex which is formed, in which a fused protein of
 CC the prey with a protein applicable as a label for detection is formed
 CC while the bait is modified for separation, and mRNA that encodes the
 CC fused protein containing the prey is translated in a cell-free

Alignment Scores:

Pred. No.:	1,66e-22	Length:	5172
Score:	274.00	Matches:	64
Percent Similarity:	63.89%	Conservative:	5
Best Local Similarity:	59.26%	Mismatches:	15
Query Match:	55.58%	Indels:	24

translation system in the presence of the bait for interaction with the produced prey before separating and detecting via the label. The method is useful for detecting interaction between substances and proteins in a cell-free translation system in vitro, which is applicable in screening interacting substances, and analysis of interactions between proteins and protein-nucleic acids as well as protein and gene functions, useful in medicine, food science and environmental monitoring. The method can be performed for global detection layerwise, even of weak interactions at high sensitivity, in the presence of a bait or prey, or full-length bait, in a cell-free translation system. But, there is no need for large-scale expression and purification of such bait by other means after production in *Escherichia coli* to greatly shorten the operating time. The present sequence represents a vector nucleotide sequence which is used in an example from the present invention.

XX SQ Sequence 5172 BP; 1365 A; 1307 C; 1291 G; 1209 T; 0 other;

Alignment Scores:

Pred. No.:	1-66e-21	Length:	5172
Score:	274.90	Matches:	64
Percent Similarity:	63.8%	Conservative:	5
Best Local Similarity:	59.2%	Mismatches:	15
Query Match:	55.5%	Indels:	24
DB:	25	Gaps:	3

US-10-059-720-19 (1-97) x ACC45014 (1-5172)

QY	11	MetAlaSerMetThrGlyGlnMetGlyArgAspProAspLeuGluGlnArgAla	30
DB	794	ATGCGTAGCATGACTGGTGGACAGCAATGGTGGCGA-----TCGGCAGAGCG	844
QY	31	GluGluLeuAlaArgGluAsn-----GluGluLeuGluGluGluGluGluGlu	48
DB	845	CAGAGCATCGGCAAGAGGGGCAAACTAGAGCAGCTATCTCTGAAGAGAGAGAAACGG	904
QY	49	GlnGluAsnAlaGlu-----	53
DB	905	AGATCCGAGGAGACGGATAGATGGCTGCGCCAGTCCCGATCGGAGGAGGAG	964
QY	54	-----LeuGluAlaGluThrAspGlnLeuGluAspGluLysSerAlaLeuGln	69
DB	965	CTGACAGATACACTCCAGCGGAGCAGATCAACTTGAAGATGAGAGTCTGCGTTCGAG	1024
QY	70	ThrGluLeuAlaAsnLeuLysGluGluGluGluGluGluGluGluGluGluGluGlu	89
DB	1025	ACTGAGATTGCCAATCTGCTGAGAGAGAGGAAACTGGAGTTTATTTGGCAGCCAC	1084

QY 90 ArgProAlaCysLysIleProAsp 97

DB 1085 CGACCTGCTGCAAGATCCCGAT 1108

RESULT 8

ABX45969

ID ABX45969 standard; cDNA; 403 BP.

XX AC ABX45969;

XX DT 21-FEB-2003 (first entry)

XX DE Bovine EST associated with lactation/muscle/fat deposition #11134.

XX KW Bovine; ss; EST; expressed sequence tag; lactation; LMFD;
XX muscle deposition; fat deposition; genome mapping; gene identification;
XX gene analysis; cattle breeding.

XX OS Bos Taurus.

XX XX US2002137139-A1.

XX PD 26-SEP-2002.

XX XX 24-SEP-2001; 2001US-0960352.

XX PF

XX XX

PR 12-JAN-1999; 99US-115707P.

PR 11-JAN-2000; 2000US-0480902.

XX (BYAT/) BYATT J C.

XX (MATH/) MATHIALAGAN N.

XX (TAON/) TAO N.

XX (WARR/) WARREN W C.

XX BYATT JC, Mathialagan N, Tao N, Warren WC;

XX WPI; 2003-110599/10.

XX New nucleic acid associated with lactation, and muscle and fat
XX deposition, useful for genome mapping, gene identification and
XX analysis, cattle breeding, or for genetically improving cattle

XX Claim 2; SEQ ID No 11134; 245pp; English.

XX The invention relates to a purified nucleic acid molecule associated with
XX lactation or muscle and fat deposition (designated LMFD), derived
XX from cattle, and the LMFD nucleic acid can specifically hybridise to a
XX second nucleic acid molecule comprising any of 1512 nucleotide
XX sequences, appearing as ABX34836-ABX49947, or complements of them.
XX Also included are; (1) a transformed cell having a nucleic acid
XX comprising an LMFD nucleic acid linked to a promoter and a 3' non-
XX translated sequence that functions in the cell to cause termination of
XX transcription and addition of polyadenylated ribonucleotides to a 3' end
XX of the mRNA molecule; and (2) determining a level or pattern of a
XX nucleic acid in a bovine cell or tissue comprising: (a) incubating a marker
XX nucleic acid (comprising any of the 1512 nucleic acid sequences or its
XX complement or fragment) with a complementary nucleic acid molecule
XX obtained from the bovine cell or tissue, where hybridisation between the
XX marker nucleic acid and the complementary nucleic acid permits the
XX detection of the molecule; and (b) detecting the level or pattern of the
XX complementary nucleic acid, where the detection of the complementary
XX nucleic acid is predictive of the level or pattern of the molecule.
XX The LMFD nucleic acid is used for determining a level or pattern
XX of a molecule in a bovine cell or tissue. It is useful for genome
XX mapping, gene identification and analysis, cattle breeding, preparation
XX of constructs for use in cattle gene expression, or for genetically
XX improving cattle. The present sequence is one of the 1512 bovine
XX LMFD EST (expressed sequence tag) nucleic acids.
XX Note: The present sequence was not shown in the specification but
XX was obtained in electronic format from the USPTO web site:
XX seqdata.uspto.gov/sequence.html?DocID=20020137139.

XX SQ Sequence 403 BP; 116 A; 113 C; 116 G; 58 T; 0 other;

Alignment Scores:

Pred. No.:	1-75e-19	Length:	403
Score:	245.00	Matches:	58
Percent Similarity:	65.35%	Conservative:	8
Best Local Similarity:	57.43%	Mismatches:	17
Query Match:	49.70%	Indels:	18
DB:	25	Gaps:	3

US-10-059-720-19 (1-97) x ABX45969 (1-403)

QY	11	MetAlaSerMetThrGlyGly-----GlnGlnMetGlyArgAspProAspLeuGluGln	28
DB	104	ATGAAACCATGATCAGAGGCGCAGAGCTCAGAGCATTCGCCCG-----AGGGGC	151
QY	29	ArgAlaGluGluLeuAlaArgGluAsnGluGluGluGluGluGluGluGluGlu	48
DB	152	AAAGGTAGAACACAGTTGTCCCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT	211
QY	49	GlnGluAsnAla-----	56
DB	212	AAAGTGGCTGCAGCCCAATGCCGAGAGGAGGAGCTGACTGACACCTCCAGCG	271
QY	57	GluThrAspGlnLeuGluAspGluLysSerAlaLeuGlnThrGluLeuAlaAsnLeu	76
DB	272	GAGACAGACCAACTAGAGATGAGAGTCCGCTTTCAGACAGAGATGTCATCTCTG	331

XX Bertucci F, Houlgate R, Birnbaum D, Nguyen C, Viens P, Fert V;
 XX WPI; 2002-619023/66.
 XX Novel polynucleotide library useful in molecular characterization of a
 XX carcinoma, comprising a pool of polynucleotide sequences or its
 XX sub-sequences which are either underexpressed or overexpressed in tumor
 XX cells -
 XX Claim 1; Page 315; 401pp; English.
 XX The present invention describes a polynucleotide library (I) useful in
 XX the molecular characterization of a carcinoma, comprising a pool of
 XX polynucleotides or its sub-sequences which are either underexpressed or
 XX overexpressed in tumor cells, and correspond to any of the
 XX polynucleotide sequences chosen from the 468 sequences given in ABV94010
 XX to ABV94477. Also described: (1) a polynucleotide array (II) useful for
 XX the prognosis or diagnostic of tumor, comprising (I); and (2) detecting
 XX (M1) differentially expressed polynucleotide sequences which are
 XX correlated with a cancer, involves obtaining a polynucleotide sample from
 XX a patient, and reacting the polynucleotide sample obtained with a probe
 XX of the polynucleotide sequences of (I) or its expression products encoded
 XX by polynucleotide sequences of (I), and detecting the reaction product.
 XX (1) have cytostatic activities and can be used as anti-tumour agents. (1)
 XX is useful in molecular characterisation of a carcinoma. (I) and (II) are
 XX useful for the prognosis or diagnostic of tumour, in differentiating a
 XX normal cell from a cancer cell, detecting a hormone sensitive tumour
 XX cell, differentiating a tumour with lymph nodes from a tumour without
 XX lymph nodes, differentiating antracycline-sensitive tumours and poor
 XX antracycline-insensitive tumours, and classifying good and poor prognosis
 XX primary breast tumours. (I) is useful for large-scale molecular
 XX characterisation of breast cancer that help in prediction, prognosis and
 XX cancer treatment, and for detecting differentially expressed genes that
 XX correlated with a cancer.
 XX SQ Sequence 2084 BP; 475 A; 580 C; 521 G; 508 T; 0 other;

Alignment Scores:
 Pred. No.: 3,566-18 Length: 2084
 Score: 241.00 Matches: 57
 Percent Similarity: 65.66% Conservative: 8
 Best Local Similarity: 57.58% Mismatches: 16
 Query Match: 48.88% Indels: 18
 DB: 24 Gaps: 3
 US-10-059-720-19 (1-97) x ABV94327 (1-2084)
 QY 13 SerMetThrGlyGly-----GlnGlnMetGlyArgAspProAspLeuGluGlnArgAla 30
 DB 495 ACCATGACAGGAGCGCGAGCGAGCGAGCATTTGGCAGG-----AGGGGCAAGGTG 542
 QY 31 GluGluLeuAlaArgGluAsnGluGluLeuGlyGluAlaGluGluGluGluGlu 50
 DB 543 GACAGNGTATCTCTCCAGAGAGAGAGAGAGAGAGAGATCCGAGGCGAAGGAATPARGATG 602
 QY 51 AsnAla-----GluLeuGluAlaGluThr 58
 DB 603 GCTGCGAGCCAAATCCGACACCGAGGAGGAGTGTACTGATACACTCCAGCGGAGACA 662
 QY 59 AspGlnLeuGluAspGluLysSerAlaLeuGlnThrGluAlaAsnLeuLeuGlu 78
 DB 663 GACCACTAGAGATGAGAGTCTGCTTTCAGAGACCGAGATTCGCAACCTCTCAAGGAG 722
 QY 79 LysGluLysLeuGluPheLeuAlaAlaHisArgProAlaCysLysIleProAsp 97
 DB 723 AAGGAAACTAGAGTTTCTCTCTGGCAGCTCACCAGCTCTCTGCAAGATCCCTGAT 779

RESULT 11
 AA241352
 ID AA241352 standard; cDNA; 2103 BP.
 XX

AA241352;
 19-JAN-2000 (first entry)
 Human normal uterus tissue derived cDNA 28.
 Human; uterus; cancer; treatment; anticancer; cytostatic; gene therapy;
 EST; expressed sequence tag; ss.
 Homo sapiens.
 DE19817946-A1.
 21-OCT-1999.
 17-APR-1998; 98DE-1017946.
 17-APR-1998; 98DE-1017946.
 (META-) METAGEN GES GENOMFORSCHUNG MBH.
 Rosenthal A, Specht T, Hinzmann B, Schmitt A, Pilarsky C, Dahl E;
 WPI; 1999-591956/51.
 New nucleic acid sequences expressed in normal uterine tissues, and
 derived polypeptides, for treatment of uterine cancer and
 identification of therapeutic agents -
 Claim 3; Page 94; 154pp; German.
 This invention describes novel cDNA sequences (A) highly expressed in
 normal uterine tissue which can have anticancer and cytostatic activity
 and can be used for gene therapy. (A) are used (i) for recombinant
 expression of polypeptides (B) and (ii) to isolate complete genes.
 (B) are used (i) to identify agents suitable for treatment of uterine
 cancer; (ii) directly for treating this form of cancer (including
 expression from gene therapy vectors) and (iii) for generation of
 specific antibodies. (A) are identified by assembling ESTs (expressed
 sequence tags) from a particular tissue type before comparison of
 expression patterns. This allows a significantly longer fragment of the
 gene to be revealed, so should reduce the number of failures associated
 with the fact that ESTs from different libraries may represent
 different parts of the same unknown gene, distorting the estimated
 frequency of occurrence in a particular tissue. AA241325-241385
 represent the human uterine tissue derived cDNA fragments of the
 invention which encode the protein fragments represented in
 AA59838-59892.
 SQ Sequence 2103 BP; 492 A; 573 C; 525 G; 513 T; 0 other;

Alignment Scores:
 Pred. No.: 3,66-18 Length: 2103
 Score: 241.00 Matches: 57
 Percent Similarity: 65.66% Conservative: 8
 Best Local Similarity: 57.58% Mismatches: 16
 Query Match: 48.88% Indels: 18
 DB: 20 Gaps: 3
 US-10-059-720-19 (1-97) x AA241352 (1-2103)
 QY 13 SerMetThrGlyGly-----GlnGlnMetGlyArgAspProAspLeuGluGlnArgAla 30
 DB 451 ACCATGACAGGAGCGCGAGCGAGCATTTGGCAGG-----AGGGGCAAGGTG 498
 QY 31 GluGluLeuAlaArgGluAsnGluGluLeuGlyGluAlaGluGluGluGluGlu 50
 DB 499 GACAGTATCTCTCCAGAGAGAGAGAGAGAGAGATCCGAGGCGAAGGAATPARGATG 558
 QY 51 AsnAla-----GluLeuGluAlaGluThr 58
 DB 559 GCTGCGAGCCAAATCCGACACCGAGGAGGAGTGTACTGATACACTCCAGCGGAGACA 618

QY 59 AspGlnLeuGluAspGluTysSerAlaLeuGlnThrGluAlaAsnLeuLeuGlu 78
 Db 619 GACCACTAGAGATGAGAGTCTGCTTTCAGACCGAGATGCACTGCTGAAGAG 678
 QY 79 LysGluLeuGluPheLeuAlaAlaHisArgProAlaCysLysIleProAsp 97
 Db 679 AAGGAAAACTAGATTCTATCTGCGCAGCTCACCAGACCTGCTGCAGATCCCTGAT 735

RESULT 12

ABQ54467
 ID ABQ54467 standard; cDNA; 2180 BP.

XX ABQ54467;

XX 22-AUG-2002 (first entry)

XX Human ovarian antigen HTHDV01 cDNA, SEQ ID NO:347.

XX Human: ovarian antigen; ovary; ovarian; breast; cancer; tumour;
 KW ovarian cancer; breast cancer; tumour; reproductive system disorder;
 KW infertility; pregnancy disorder; anovulation; polycystic ovary syndrome;
 KW PCOS; ovarian cyst; dysmenorrhea; endocrine disorder; infection;
 KW inflammatory condition; immune disorder; blood disorder;
 KW cardiovascular disorder; respiratory disorder; neurological disorder;
 KW gastrointestinal disorder; urinary system disorder; drug screening;
 KW gene therapy; chromosome mapping; forensic analysis;
 KW antibody preparation; cytostatic; immunomodulatory; neuroprotective;
 KW antiinflammatory; gynaecological; reproductive; chromosome 14q24.3;
 KW gene, ss.

XX Homo sapiens.

XX WO200200677-A1.

XX 03-JAN-2002.

XX 07-JUN-2001; 2001WO-US18569.

XX 07-JUN-2000; 2000US-209467P.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Birse CE, Rosen CA;

XX MPI; 2002-147878/19.

XX P-PSDB; ABP41390.

XX Isolated nucleic acid molecules encoding novel ovarian polypeptides,
 PT useful in the prevention, treatment and diagnosis of cancer (e.g.
 PT ovarian cancer), immune disorders, cardiovascular disorders and
 PT neurological diseases -

XX Claim 1; SEQ ID NO 347; 2922pp; English.

XX The invention relates to 2175 novel human ovarian antigens (ABP41054-
 CC ABP43228) and to cDNAs encoding them (ABQ54131-ABQ56305), and also
 CC encompasses polypeptides 90% identical and polynucleotides 95% identical
 CC to the sequences of the invention. The invention additionally relates to
 CC recombinant vectors and host cells comprising human ovarian antigen
 CC polynucleotides, antibodies against human ovarian antigens, and the use
 CC of ovarian antigen polynucleotides and polypeptides in diagnosing,
 CC treating, prognosing or preventing various ovary and/or breast-related
 CC disorders. Such conditions include ovarian cancer and breast cancer, and
 CC metastatic tumours of ovarian or breast origin, reproductive system
 CC disorders (e.g., infertility, disorders of pregnancy, anovulation,
 CC polycystic ovary syndrome, ovarian cysts, and dysmenorrhea), endocrine
 CC disorders, infections (e.g., chlamydia, HIV, toxoplasmosis, and toxic
 CC shock syndrome), inflammatory conditions (e.g., mastitis, oophoritis and
 CC vaginitis), immune disorders (e.g., congenital and acquired
 CC immunodeficiencies, autoimmune oophoritis, systemic lupus erythematosus),
 CC blood-related disorders (e.g., anaemia), cardiovascular disorders,
 CC respiratory disorders, neurological disorders, gastrointestinal disorders
 CC and urinary system disorders. Ovarian antigen polypeptides and

CC polynucleotides may also be used in screening for compounds which
 CC modulate ovarian antigen expression or activity. The polynucleotides may
 CC further be used for gene therapy, chromosome mapping, in the
 CC identification of individuals and in forensic analysis, and the
 CC polypeptides may be used as food additives or to prepare antibodies
 CC useful in disease diagnosis, drug targeting and phenotyping. The present
 CC sequence represents cDNA encoding a human ovarian antigen of the
 CC invention.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 2180 BP; 513 A; 596 C; 549 G; 519 T; 3 other;

Alignment Scores:

Pred. No.: 3.76e-18 Length: 2180
 Score: 241.00 Matches: 57
 Percent Similarity: 65.66% Conservative: 8
 Best Local Similarity: 57.58% Mismatches: 16
 Query Match: 48.88% Indels: 18
 DB: 24 Gaps: 3

US-10-059-720-19 (1-97) x ABQ54467 (1-2180)

QY 13 SerMetThrGlyGly-----GlnGlnMetGlyArgAspProAspLeuGluGlnArgAla 30
 Db 492 ACCATGACAGGAGGCGGAGCGCAGAGCATTCGCAGG-----ACGGGCAAGGTG 539
 QY 31 GluGluLeuAlaArgGluAsnGluGluGluGlyGluAlaGluGluGluGluGlu 50
 Db 540 GAACAGTTATCTCCAGAGAGAGAGAGAAAGAGGAATCCAGAGGGAAGGAATAAGATG 599
 QY 51 AsnAla-----GluLeuGluAlaGluThr 58
 Db 600 GCTGCAGCAAAATGCGCAACCGGAGGAGGAGCTGACTGATACACTCCAGCGGAGACA 659
 QY 59 AspGlnLeuGluAspGluIysSerAlaLeuGlnThrGluIleAlaAsnLeuLeuGlu 78
 Db 660 GACCACTAGAGATGAGAGTCTGCTTTCAGACCGAGATGCTCAACCTGCTGAAGGAG 719
 QY 79 LysGluLeuGluPheIleLeuAlaAlaHisArgProAlaCysLysIleProAsp 97
 Db 720 AAGGAAAACTAGATTCTATCTGCGCAGCTCACCAGACCTGCTGCAGATCCCTGAT 776

RESULT 13

AA184355
 ID AA184355 standard; DNA; 296 BP.

XX AA184355;

XX 12-NOV-1997 (first entry)

XX 4Heptad-P zipper DNA sequence.

XX DNA binding protein; RNA binding protein; amphipathic peptide;
 KW acidic extension peptide; gene control; gene regulation;
 KW transcription; dominant negative protein; cancer; drug therapy;
 KW drug design; ss.

XX Synthetic.

XX Key Location/Qualifiers
 FT CDS 1..285 /*tag= a
 FT sig_peptide 1..42 /*tag= b

XX WO9705249-A2.

XX 13-FEB-1997.

XX 31-JUL-1996; 96WO-US12590.

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: November 16, 2003, 01:07:06 ; Search time 2858 Seconds
(without alignments)
1388.465 Million cell updates/sec

Title: US-10-059-720-19

Perfect score: 493

Sequence: 1 MDYDDDDKHMASNTGGQOM.....EKEKLEFLAHRPACKIPD 97

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2888711 seqs, 2045481386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-O=/cgn2.1/USPTO.spool/US10059720/runat.14112003.184959.17887/app.query.fasta.1.263
-DB=genemb1 -QFMT=fastcap -SUFFIX=xge -MINMATCH=0.1 -LCOPCL=0 -LCOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOALIGN=200 -THR SCORE=500 -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10059720.scgn.1.1.3508 @runat.14112003.184959.17887 -NCPU=6 -ICPU=3
-NO MMAP -LARGQUERY -NEG SCORE=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -PGAPOP=6
-XGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl.*
1: gb.ba.*
2: gb.hcg.*
3: gb.in.*
4: gb.om.*
5: gb.ov.*
6: gb.pat.*
7: gb.ph.*
8: gb.pl.*
9: gb.pr.*
10: gb.ro.*
11: gb.sts.*
12: gb.sy.*
13: gb.un.*
14: gb.vi.*
15: em.ba.*
16: em.fun.*
17: em.hum.*
18: em.in.*
19: em.mu.*
20: em.om.*
21: em.or.*
22: em.ov.*
23: em.pat.*
24: em.ph.*
25: em.pl.*
26: em.ro.*
27: em.sts.*
28: em.un.*

29: em.vi.*
30: em.htg.hum.*
31: em.htg.inv.*
32: em.htg.other.*
33: em.htg.mus.*
34: em.htg.pln.*
35: em.htg.rtd.*
36: em.htg.mam.*
37: em.htg.vrt.*
38: em.sy.*
39: em.htgo.hum.*
40: em.htgo.mus.*
41: em.htgo.other.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	493	100.0	300	6	AR202098 Sequence
2	441	89.5	281	6	AR202097 Sequence
3	350	71.0	324	6	AR202096 Sequence
4	288	58.4	266	6	AR202094 Sequence
5	274	55.6	418	6	BD169720 C-terminu
6	241	48.9	1823	9	BC004490 Homo sapi
7	241	48.9	1918	9	AK097379 Homo sapi
8	241	48.9	2084	6	AX587848 Sequence
9	241	48.9	2103	6	AX014320 Sequence
10	238	48.3	296	6	AR202105 Sequence
11	233	47.3	2107	10	BC029814 Mus muscu
12	232	47.1	2116	10	RNCFOSR Rat c-fos m
13	231	46.9	1146	10	AF033012 Cricetulu
14	231	46.9	3811	14	AF033814 Murine os
15	231	46.9	3811	14	MSVMUSV KO2712 PER murine
16	230.5	46.8	2015	4	AF540379 Felis cat
17	229	46.5	262	6	AR202095 Sequence
18	229	46.5	4226	14	REMSV5 V01184 Provirus of
19	224	45.4	297	10	RATWBFOS M34001 Rat wt-Fos
20	219	44.4	1104	5	AF364329 Coturnix
21	219	44.4	1980	14	AC2NK24 M18041 Avian trans
22	219	44.4	2253	5	CHKCFOSA M37000 Chicken c-f
23	219	44.4	2253	5	GDFOS Y00659 Chicken fos
24	216	43.8	228	4	OR553325 Ovis arie
25	216	43.8	273	6	AX098479 Sequence
26	216	43.8	381	4	AF069515 Bos tauru
27	210.5	42.7	200387	2	AC140958 Delphis
28	207	42.0	1686	5	CCU81505 U81505 Cyprinus ca
29	207	42.0	6023	10	AF061881 Mesocric
30	206	41.8	414	4	OAY15747 Y15747 Ovis arie
31	206	41.8	414	4	SSCFQ524 Y14808 Sus scrofa
32	206	41.8	700	9	AB022276 Homo sapi
33	206	41.8	3565	6	AX330284 Sequence
34	206	41.8	3565	6	AX336394 Sequence
35	206	41.8	3565	6	AX663638 Sequence
36	206	41.8	3565	6	I96207 Sequence 44
37	206	41.8	3565	9	HSCFOS V01512 Human cellu
38	206	41.8	3565	10	MUSFOS J00370 Mouse c-fos
39	206	41.8	3967	6	AX306224 Sequence
40	206	41.8	3967	6	BD078052 JNK3-cont
41	206	41.8	3967	10	V00727 Mouse c-fos
42	206	41.8	4200	4	SSC132510 A132510 Sus scrof
43	206	41.8	5897	9	AY212879 Homo sapi
44	206	41.8	6210	6	AX663628 Sequence
45	206	41.8	6210	9	HUMFOS X00650 Human fos p

ALIGNMENTS

RESULT 1

AR202098
LOCUS AR202098 300 bp DNA linear PAT 20-APR-2002
DEFINITION Sequence 18 from patent US 6361968.
ACCESSION AR202098
VERSION AR202098.1 GI:20256637
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 300)
AUTHORS Vinson,C.R. and Krylov,D.
TITLE Extension of a protein-protein interaction surface to inactive the function of a cellular protein
JOURNAL Patent: US 6361968-A 18 26-MAR-2002;
FEATURES Location/Qualifiers
source 1..300
BASE COUNT 96 a 68 c 89 g 47 t
ORIGIN
Alignment Scores:
Pred. No.: 6,76e-39 Length: 300
Score: 493.00 Matches: 97
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0
US-10-059-720-19 (1-97) x AR202098 (1-300)
QY 1 MetAspTyrLysAspAspAspLysHisMetAlaSerMetThrGlyGlnGlnMet 20
Db 3 ATGGACTACAGGAGGAGATGACAGCATATGGCTAGCATGCTGGTGGACCAATG 62
QY 21 GlyArgAspProAspLeuGlnArgAlaGluGluLeuAlaArgGluAsnGluGluLeu 40
Db 63 GGTGGGATCCTGACCTGCAACAGCTGCTGAGGAACTGGCCGCGGAGAAAGAGAGCTG 122
QY 41 GluLysGluAlaGluGluLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 60
Db 123 GAAAGAGGCGGAGAGCTGGAGCAGGAGAAACGCTGAACCTGAGCGGAGACACCAA 182
QY 61 LeuGluAspGluLysSerAlaLeuGlnThrGluLeuAlaAsnLeuLysGluLysGlu 80
Db 183 CTAGAGATGAGAGCTGCTTTGAGACCGAGATGCGAACCTGCTGAGGAGAGGAA 242
QY 81 LysLeuGluPheLeuAlaHisArgProAlaCysLysLysIleProAsp 97
Db 243 AAACCTAGAGTTTCCTGGCAGCTCAGCGACCTGCTGCAAGATCCCTGAT 293
RESULT 2
AR202097
LOCUS AR202097 281 bp DNA linear PAT 20-APR-2002
DEFINITION Sequence 16 from patent US 6361968.
ACCESSION AR202097
VERSION AR202097.1 GI:20256636
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 281)
AUTHORS Vinson,C.R. and Krylov,D.
TITLE Extension of a protein-protein interaction surface to inactive the function of a cellular protein
JOURNAL Patent: US 6361968-A 16 28-MAR-2002;
FEATURES Location/Qualifiers
source 1..281
BASE COUNT 90 a 62 c 81 g 48 t
ORIGIN
Alignment Scores:
Pred. No.: 5,73e-34 Length: 281
Score: 493.00 Matches: 97
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0
US-10-059-720-19 (1-97) x AR202097 (1-281)
QY 10 HisMetAlaSerMetThrGlyGlnGlnMetGlyArgAspProAspLeuGlnArg 29
Db 6 CATATGGCTAGCATGACTGGTGGAGCAAAATGGTGGGATCTGCTGACCTGGAAACAAGT 65
QY 30 AlaGluGluLeuAlaArgGluAsnGluGluLeuGluLysGluAlaGluGluGln 49
Db 66 GCTGAGGAACCTGGCCCGTGAACAGAGAGCTGGAAAGAGGCGCGAGAGCTGGAGCAG 125
QY 50 GluAsnAlaGluLeuGluAlaGluThrAspGlnLeuGluAspGluLysSerAlaLeuGln 69
Db 126 GAAACCGCTGAACCTGAGCGGAGACAGCAACTAGAAAGATGAGAGTCTGCTTTGCAG 185
QY 70 ThrGluLeuAlaAsnLeuLysGluLysGluLysGluLysGluLysGluLysGluLys 89
Db 186 ACCGAGATTCACACCTGCTGAGGAGAGGAGAAACTAGAGTTTCATCTGGCAGCTCAC 245
QY 90 ArgProAlaCysLysLysIleProAsp 97
Db 246 CGACCTGCTGCAAGATCCCTGAT 269
RESULT 3
AR202096
LOCUS AR202096 324 bp DNA linear PAT 20-APR-2002
DEFINITION Sequence 14 from patent US 6361968.
ACCESSION AR202096
VERSION AR202096.1 GI:20256635
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 324)
AUTHORS Vinson,C.R. and Krylov,D.
TITLE Extension of a protein-protein interaction surface to inactive the function of a cellular protein
JOURNAL Patent: US 6361968-A 14 26-MAR-2002;
FEATURES Location/Qualifiers
source 1..324
BASE COUNT 109 a 71 c 93 g 51 t
ORIGIN
Alignment Scores:
Pred. No.: 3,16e-25 Length: 324
Score: 350.00 Matches: 72
Percent Similarity: 74.53% Conservative: 7
Best Local Similarity: 67.92% Mismatches: 17
Query Match: 70.99% Indels: 10
DB: 6 Gaps: 2
US-10-059-720-19 (1-97) x AR202096 (1-324)
QY 1 MetAspTyrLysAspAspAspLysHisMetAlaSerMetThrGlyGlnGlnMet 20
Db 3 ATGGACTACAGGAGGAGATGACAGCATATGGCTAGCATGCTGGTGGACAGCAATG 62
QY 21 GlyArgAspProAspLeuGlnArgAlaGluGluLeuAlaArgGluAsnGluGluLeu 40
Db 63 GGTGGGATCCCAAGTGGAGACAGATTCCTCCAGAA---GAAGAGAGAAAGAGAGATC 119
QY 41 GluLysGluAlaGluGluLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 53
Db 120 CGAAGGGAAGGAATAAGATGGCTGAGCAAAATGCGCAACCGGAGGAGGAGCTGACT 179
QY 54 -----LeuGluAlaGluThrAspGlnLeuGluAspGluLysSerAlaLeuGlnThrGlu 71
Db 119

Score: 441.00 Matches: 88
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 89.45% Indels: 0
DB: 6 Gaps: 0
US-10-059-720-19 (1-97) x AR202097 (1-281)
QY 10 HisMetAlaSerMetThrGlyGlnGlnMetGlyArgAspProAspLeuGlnArg 29
Db 6 CATATGGCTAGCATGACTGGTGGAGCAAAATGGTGGGATCTGCTGACCTGGAAACAAGT 65
QY 30 AlaGluGluLeuAlaArgGluAsnGluGluLeuGluLysGluAlaGluGluGln 49
Db 66 GCTGAGGAACCTGGCCCGTGAACAGAGAGCTGGAAAGAGGCGCGAGAGCTGGAGCAG 125
QY 50 GluAsnAlaGluLeuGluAlaGluThrAspGlnLeuGluAspGluLysSerAlaLeuGln 69
Db 126 GAAACCGCTGAACCTGAGCGGAGACAGCAACTAGAAAGATGAGAGTCTGCTTTGCAG 185
QY 70 ThrGluLeuAlaAsnLeuLysGluLysGluLysGluLysGluLysGluLysGluLys 89
Db 186 ACCGAGATTCACACCTGCTGAGGAGAGGAGAAACTAGAGTTTCATCTGGCAGCTCAC 245
QY 90 ArgProAlaCysLysLysIleProAsp 97
Db 246 CGACCTGCTGCAAGATCCCTGAT 269
RESULT 3
AR202096
LOCUS AR202096 324 bp DNA linear PAT 20-APR-2002
DEFINITION Sequence 14 from patent US 6361968.
ACCESSION AR202096
VERSION AR202096.1 GI:20256635
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 324)
AUTHORS Vinson,C.R. and Krylov,D.
TITLE Extension of a protein-protein interaction surface to inactive the function of a cellular protein
JOURNAL Patent: US 6361968-A 14 26-MAR-2002;
FEATURES Location/Qualifiers
source 1..324
BASE COUNT 109 a 71 c 93 g 51 t
ORIGIN
Alignment Scores:
Pred. No.: 3,16e-25 Length: 324
Score: 350.00 Matches: 72
Percent Similarity: 74.53% Conservative: 7
Best Local Similarity: 67.92% Mismatches: 17
Query Match: 70.99% Indels: 10
DB: 6 Gaps: 2
US-10-059-720-19 (1-97) x AR202096 (1-324)
QY 1 MetAspTyrLysAspAspAspLysHisMetAlaSerMetThrGlyGlnGlnMet 20
Db 3 ATGGACTACAGGAGGAGATGACAGCATATGGCTAGCATGCTGGTGGACAGCAATG 62
QY 21 GlyArgAspProAspLeuGlnArgAlaGluGluLeuAlaArgGluAsnGluGluLeu 40
Db 63 GGTGGGATCCCAAGTGGAGACAGATTCCTCCAGAA---GAAGAGAGAAAGAGAGATC 119
QY 41 GluLysGluAlaGluGluLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 53
Db 120 CGAAGGGAAGGAATAAGATGGCTGAGCAAAATGCGCAACCGGAGGAGGAGCTGACT 179
QY 54 -----LeuGluAlaGluThrAspGlnLeuGluAspGluLysSerAlaLeuGlnThrGlu 71
Db 119

AUTHORS TITLE JOURNAL

Strausberg, R.
Direct Submission
Submitted (12-MAR-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA

REMARK COMMENT

NIH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: cgapbs-remail.nih.gov
Tissue procurement: ATCC
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Institute for Systems Biology
<http://www.systemsbiology.org>
contact: amad@systemsbiology.org
Anup Madan, Rachel Dickhoff, Jessica Fahey, Stephanie Ford, Julia
Greene, Mark Ketterman and Anuradha Madan

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IPAL Plate: 14 Row: O Column: 17
This clone was selected for full length sequencing because it
passed the following selection criteria: Hexamer frequency ORF
analysis, similarity but not identity to protein.

FEATURES source

1..1823
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="LocusID:2353"
/db_xref="taxon:9606"
/clone="MGC:11074 IMAGE:3688670"
/tissue_type="pancreas, adenocarcinoma"
/clone_lib="NIH MGC 39"
/lab_host="DH10B-R"
/note="Vector: pOTB7"
136..1278
/codon_start=1
/product="v-fos FBJ murine osteosarcoma viral oncogene
homolog"
/protein_id="AAH04490.1"
/db_xref="GI:13325364"
/translation="MFEFSGNADYEASSRCSASPAGDSLSVYHSPADSFSSMGSPV
NAQDFCTDLAVSSANFTPTVAISTSPDLQVLPALVSSVAPSCQTRAPFGVPAS
AGAYSRAGVVTWGTGGAQSGIRGKVEQLSPEERGRIRKRNKAAKRNRRRE
LDTQATDQLEKALQTEIANLKEKLEFLAAHRPAKIPDDIGFPEEMSV
ASDLTGLEPEVATPESEEAFTPLNDPEKPSVFPVKSISSMELTEPFDPLFPA
SSRPSGSETARSPMDLSGFSYAADMEPLHSGSLGMPWATELEPLCTPVVTCPTSC
TAYTSSFFVTPEADSPFSCAAHRKGSNSNPSDSLSLSPFLAL"

CDS

416 a 545 c 476 g 386 t
/product="v-fos FBJ murine osteosarcoma viral oncogene
homolog"
/protein_id="AAH04490.1"
/db_xref="GI:13325364"
/translation="MFEFSGNADYEASSRCSASPAGDSLSVYHSPADSFSSMGSPV
NAQDFCTDLAVSSANFTPTVAISTSPDLQVLPALVSSVAPSCQTRAPFGVPAS
AGAYSRAGVVTWGTGGAQSGIRGKVEQLSPEERGRIRKRNKAAKRNRRRE
LDTQATDQLEKALQTEIANLKEKLEFLAAHRPAKIPDDIGFPEEMSV
ASDLTGLEPEVATPESEEAFTPLNDPEKPSVFPVKSISSMELTEPFDPLFPA
SSRPSGSETARSPMDLSGFSYAADMEPLHSGSLGMPWATELEPLCTPVVTCPTSC
TAYTSSFFVTPEADSPFSCAAHRKGSNSNPSDSLSLSPFLAL"

BASE COUNT

416 a 545 c 476 g 386 t
Alignment Scores:
Pred. No.: 4.75e-14 Length: 1823
Score: 241.00 Matches: 57
Percent Similarity: 65.66% Conservative: 8
Best Local Similarity: 57.58% Mismatches: 16
Query Match: 48.88% Indels: 18
DB: 9 Gaps: 3

US-10-059-720-19 (1-97) x BC004490 (1-1823)

Qy 13 SerMetThrGlyGly-----GlnGlnMetGlyArgAspProAspLeuGluGlnArgAla 30
Db 475 ACCATGACAGAGAGCGCCGAGCGCAGCATTCGCGAGG-----AGGGGCAAGGTG 522
Qy 31 GluGluLeuAlaArgGluAsnGluGluLeuGluGluGluGluGluGluGluGluGlu 50
Db 523 GAACAGTATCTCCAG 582
Qy 51 AsnAla-----GluLeuGluAlaGluThr 58
Db 583 GCTGACGACCAAAATCCCGCAACCCGAGGAGGAGGAGTGTATGATCACTCCAAGCGGAGACA 642

Qy

59 AspGlnLeuGluAspGluLysSerAlaLeuGlnThrGlnLeuAlaAsnLeuLeuGlu 78
Db 643 GACCACTAGAGATGAGAGATCTGCTTTCGAGCCGAGATTCACCACTGCTCGAGAG 702

Qy

79 LysGluLysLeuGluPheLeuAlaAlaHisArgProAlaCysLysLysLeuProAsp 97
Db 703 AAGGAAAACATAGAGTTCATCTGCGACGCTCACCAGACCTGCTGCAAGATCCCTGAT 759

RESULT 7

AK097379 1918 bp mRNA linear PRI 15-JUL-2002
Homo sapiens cDNA FLJ40060 fis, clone TC0LN2000236, highly similar
to P55-C-FOS PROTO-ONCOGENE PROTEIN.
AK097379
VERSION
KEYWORDS
SOURCE
ORGANISM

REFERENCE

1
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
Ninomiya, K., Wagatsuma, M., Kanda, K., Kondo, H., Yokoi, T.,
Kodaira, H., Furuya, T., Takahashi, M., Kikkawa, E., Omura, Y., Abe, K.,
Kaminara, K., Katsuta, N., Sato, K., Tanikawa, M., Yamazaki, M.,
Sugiyama, T., Irie, R., Otsuki, T., Sato, H., Wakamatsu, A., Ishii, S.,
Yamamoto, J., Isono, Y., Kawai-Hio, Y., Saito, K., Nishikawa, T.,
Kimura, K., Yamashita, H., Matsuoka, K., Nakamura, Y., Sekine, M.,
Kikuchi, H., Murakawa, K., Kanehori, K., Takahashi-Fujii, A.,
Oshima, A., Sugiyama, A., Kawakami, B., Suzuki, Y., Sugano, S.,
Nagahara, K., Masuho, Y., Nagai, K. and Isogai, T.
Unpublished
NEDO human cDNA sequencing project

AUTHORS

Isogai, T. and Yamamoto, J.
Direct Submission
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL

COMMENT

Submitted (04-JUL-2002) Takao Isogai, FLJ Project (HRI Team); 2-6-7
Kazusa-Kamatari, Kisarazu, Chiba 292-0812, Japan
(E-mail: genomics@hri.co.jp, Tel: 81-438-52-3975, Fax: 81-438-52-3986)
NEDO human cDNA sequencing project supported by Ministry of
Economy, Trade and Industry of Japan; cDNA full insert sequencing:
Research Association for Biotechnology (RAB); cDNA library
construction: Helix Research Institute (HRI) (supported by Japan
Key Technology Center etc.); 5' - 3' - end one pass sequencing: RAB,
HRI, and Biotechnology Center, National Institute of Technology and
Evaluation; clone selection for full insert sequencing: HRI and
RAB; annotation: HRI and RAB.

FEATURES

Location/Qualifiers
1..1918
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="TC0LN2000236"
/tissue_type="colon, tumor tissue"
/clone_lib="TC0LN2"
/notes="cloning vector: pME18SFL3"
BASE COUNT 452 a 505 c 482 g 479 t
ORIGIN

Alignment Scores:

Pred. No.: 5.01e-14 Length: 1918
Score: 241.00 Matches: 57
Percent Similarity: 65.66% Conservative: 8
Best Local Similarity: 57.58% Mismatches: 16
Query Match: 48.88% Indels: 18
DB: 9 Gaps: 3

US-10-059-720-19 (1-97) x AK097379 (1-1918)

Qy 13 SerMetThrGlyGly-----GlnGlnMetGlyArgAspProAspLeuGluGlnArgAla 30
Db 310 ACCATGACAGAGAGCGCCGAGCGCAGCATTCGCGAGG-----AGGGGCAAGGTG 357
Qy 31 GluGluLeuAlaArgGluAsnGluGluGluGluGluGluGluGluGluGluGluGlu 50

```

Db 358 GACAGTATCTCTGACAGAGAGAGAAAGGAGATCCGAGGGAAGCAATAGATG 417
Qy 51 AsnAla-----GluLeuGluAlaGluThr 58
Db 418 GCTGACGCAAAATGCCGCAACGGAGGAGGAGCTGACTGATACACTCCAAAGCGGAGACA 477
Qy 59 AspGlnLeuGluAspGluLeuSerAlaLeuGlnThrGluIleAlaAsnLeuLeuysGlu 78
Db 478 GACCACTAGAGATGAGAGATCTGCTTTCGACACCGAGATTGCAACCTGCTGAGGAG 537
Qy 79 LysGluLysLeuGluPheIleLeuAlaAlaHisArgProAlaCysLysIleProAsp 97
Db 538 AAGGAAAAAAGTACAGTTTCATCTGCGAGCTCACCAGCTGCTGCTCAAGATCCCTGAT 594

RESULT 8
LOCUS AX587848 2084 bp DNA linear PAT 10-JAN-2003
DEFINITION Sequence 318 from Patent WO0246467.
ACCESSION AX587848
VERSION AX587848.1 GI:28212432
KEYWORDS
SOURCE synthetic construct
ORGANISM synthetic construct
REFERENCE
AUTHORS Bertucci,F., Houligatte,R., Birnbaum,D., Nguyen,C., Viens,P. and
Fert,V.
TITLE Gene expression profiling of primary breast carcinomas using arrays
of candidate genes
JOURNAL Patent: WO 0246467-A 318 13-JUN-2002;
Ipsogen (FR)
FEATURES
source
1..2084
Location/Qualifiers
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="primer"
misc_feature 1..2084
/note="v-fos fbj murine osteosarcoma viral oncogene
homolog (FOS) gene."
BASE COUNT 475 a 580 c 521 g 508 t
ORIGIN
Alignment Scores:
Pred. No.: 5,468-14 Length: 2084
Score: 241.00 Matches: 57
Percent Similarity: 65.66% Conservative: 8
Best Local Similarity: 57.58% Mismatches: 16
Query Match: 48.88% Indels: 18
DB: Gaps: 3
US-10-059-720-19 (1-97) x AX587848 (1-2084)
Qy 13 SerMetThrGlyGly-----GlnGlnMetGlyArgAspProAspLeuGluGlnArgAla 30
Db 495 ACCATGACAGGAGGCGGAGCGAGAGCATTCGACAGG-----AGGGGCAAGGTG 542
Qy 31 GluGluLeuAlaArgGluAsnGluLeuGluLysGluAlaGluGluGlnGlu 50
Db 543 GACAGATTATCTCCAGAGAGAGAGAAAGGAGATCCGAGGGAAGCAATAGATG 602
Qy 51 AsnAla-----GluLeuGluAlaGluThr 58
Db 603 GCTGACGCAAAATGCCGCAACGGAGGAGGAGCTGACTGATACACTCCAAAGCGGAGACA 662
Qy 59 AspGlnLeuGluAspGluLeuSerAlaLeuGlnThrGluIleAlaAsnLeuLeuysGlu 78
Db 663 GACCACTAGAGATGAGAGATCTGCTTTCGACACCGAGATTGCAACCTGCTGAGGAG 722
Qy 79 LysGluLysLeuGluPheIleLeuAlaAlaHisArgProAlaCysLysIleProAsp 97
Db 723 AAGGAAAAAAGTACAGTTTCATCTGCGAGCTCACCAGCTGCTGCTCAAGATCCCTGAT 779

```

```

RESULT 9
LOCUS AX014320 2103 bp DNA linear PAT 07-SEP-2000
DEFINITION Sequence 28 from Patent WO9954353.
ACCESSION AX014320
VERSION AX014320.1 GI:10040674
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE
AUTHORS Schmitt,A., Specht,T., Dahl,E., Hinzmänn,B., Rosenthal,A. and
Pilarczyk,C.
TITLE Human nucleic acid sequences of normal uterus tissue
JOURNAL Patent: WO 9954353-A 28 28-OCT-1999;
SCHMITT ARMIN (DE); SPECHT THOMAS (DE); DAHL EDGAR (DE); HINZMANN
BERND (DE); ROSENTHAL ANDRE (DE); METAGEN GES FUER GENOMFORSCHUN
(DE); PILARSKY CHRISTIAN (DE)
FEATURES
source
1..2103
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
BASE COUNT 492 a 573 c 525 g 513 t
ORIGIN
Alignment Scores:
Pred. No.: 5,528-14 Length: 2103
Score: 241.00 Matches: 57
Percent Similarity: 65.66% Conservative: 8
Best Local Similarity: 57.58% Mismatches: 16
Query Match: 48.88% Indels: 18
DB: Gaps: 3
US-10-059-720-19 (1-97) x AX014320 (1-2103)
Qy 13 SerMetThrGlyGly-----GlnGlnMetGlyArgAspProAspLeuGluGlnArgAla 30
Db 451 ACCATGACAGGAGGCGGAGCGAGAGCATTCGACAGG-----AGGGGCAAGGTG 498
Qy 31 GluGluLeuAlaArgGluAsnGluLeuGluLysGluAlaGluGluGlnGlu 50
Db 499 GACAGATTATCTCCAGAGAGAGAGAAAGGAGATCCGAGGGAAGCAATAGATG 558
Qy 51 AsnAla-----GluLeuGluAlaGluThr 58
Db 559 GCTGACGCAAAATGCCGCAACGGAGGAGGAGCTGACTGATACACTCCAAAGCGGAGACA 618
Qy 59 AspGlnLeuGluAspGluLysSerAlaLeuGlnThrGluIleAlaAsnLeuLeuysGlu 78
Db 619 GACCACTAGAGATGAGAGATCTGCTTTCGACACCGAGATTGCAACCTGCTGAGGAG 678
Qy 79 LysGluLysLeuGluPheIleLeuAlaAlaHisArgProAlaCysLysIleProAsp 97
Db 679 AAGGAAAAAAGTACAGTTTCATCTGCGAGCTCACCAGCTGCTGCTCAAGATCCCTGAT 735

RESULT 10
LOCUS AR202105 296 bp DNA linear PAT 20-APR-2002
DEFINITION Sequence 32 from patent US 6361968.
ACCESSION AR202105
VERSION AR202105.1 GI:20256644
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE
AUTHORS Vinson,C.R. and Krylov,D.
TITLE Extension of a protein-protein interaction surface to inactive the
function of a cellular protein
JOURNAL Patent: US 6361968-A 32 26-MAR-2002;
FEATURES
Location/Qualifiers

```

source 1. .296
BASE COUNT 78 a 64 c 106 g 48 t
ORIGIN /organism="unknown"

Alignment Scores:
Pred. No.: 1.37e-14 Length: 296
Score: 238.00 Matches: 53
Percent Similarity: 75.32% Conservative: 5
Best Local Similarity: 68.83% Mismatches: 17
Query Match: 48.28% Indels: 2
DB: 6 Gaps: 1

US-10-059-720-19 (1-97) x AR202105 (1-296)

QY 11 MetAlaSerMetThrGlyGlnMetGlyArgAspPro-----AspLeuGluGln 28
Db 1 ATGGCTAGCATGACTGGTGGACAGCAATGGTCGGGATCTCGCGTGGCGCTGGAAACA 60
QY 29 ArgAlaGluLeuAlaArgGluAsnGluLeuGluLysGluAlaGluGluGlu 48
Db 61 CGTGCTAGGAACTGGCCCGTGGAAACGACAGCTGGAAAGAGCCGAGAGCTGGAG 120
QY 49 GlnGluAlaGluLeuGluAlaGluThrAspGlnLeuGluAspGluLysSerAlaLeu 68
Db 121 CAGGAAACGCTGAACCTGACGAGAGTGTGGAGTTGGAAGTGGAAAGCTGAATGACCGCTG 180
QY 69 GlnThrGluLeuAlaAsnLeuLysGluLysGluLysGluGluPheLe 85
Db 181 CGCAAGGAGTGGACAGCTGGAGCTGAGTGGAGTGGACAGCTGGCGGGTATC 231

RESULT 11
LOCUS BC029814
DEFINITION MUS MUSCULUS FBJ OSTEOSARCOMA ONCOGENE, mRNA (CDNA clone MGC:36690
IMAGE:2582234), complete cds.
ACCESSION BC029814
VERSION BC029814.1 GI:20988256
KEYWORDS MGC.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 2107)
Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,
Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,
Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,
Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F.,
Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,
Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,
Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S.,
Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J.,
Abramson, R.D., Mullahy, S.J., Bosak, S.A., McEwan, P.J.,
McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,
Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,
Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
Fahney, J., Heltin, E., Kettman, M., Madan, A., Rodriguez, S.,
Sanchez, A., Whitting, M., Madan, A., Young, A.C., Shevchenko, Y.,
Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,
Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,
Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smalios, D.E.,
Schnerch, A., Schein, J.E., Jones, S.J., and Marra, M.A.
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
2388257
PUBMED 12477932
2 (bases 1 to 2107)
Strausberg, R.
Direct Submission
Submitted (06-MAY-2002) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,

USA
NIH-MGC Project URL: http://mgc.nci.nih.gov
Contact: MGC help desk
Email: cgapbs-remail.nih.gov
Tissue Procurement: Jeffrey Green M.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Baylor College of Medicine Human Genome
Sequencing Center
Center code: BCM-HGSC
Web site: http://www.hgsc.bcm.tmc.edu/cdna/
Contact: amg@bcm.tmc.edu
Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Lounseged, H.,
Kowis, C.R., Sneed, A.J., Martin, R.G., Muzny, D.M., Nanavati,
A.N., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAC Plate: 61 Row: a Column: 1.

FEATURES
source
1. .2107
/organism="Mus musculus"
/mol_type="mRNA"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="MGC:36690 IMAGE:2582234"
/tissue_type="Mammary tumor; C3(1)-Tag model. Infiltrating
ductal carcinoma. 5 month old virgin mouse."
/clone_lib="NCI CGAP_Mam6"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6"
1. .2107
/gene="Fos"
/note="synonym: c-fos"
/db_xref="LocusID:14281"
/db_xref="MGI:95574"
140. .1282
/product="FBJ osteosarcoma oncogene"
/protein_id="AAH29814.1"
/db_xref="GI:20988257"
/db_xref="LocusID:14281"
/translation="MMFSGFNADYEASRSCSSASGASGASLSVYHSPADSPSSMGSPV
NTQDFCADLSSVSNAPITPTVTAISTSPDLQVLQVPTLVSSVAPSTQPRAPHYCLPTQS
AGYAPAGMVKTVSGRAGISGRGKVEQLSPSEKRIIRERKNKAAKKNRRE
LDYTLAQDQLEDEKSAQTIANLKEKLEFLIAHRAKPKIPDGLGFEENSV
ASDLTGGLPEASTPESEAFLLPLNDPEPLPSLEPVKSIINVELKAPFFDDFLPPA
SSRPSGSETSRSPVDYLSGSFYAADWEPLHNSLGMGMVTELEPLCTPVVVTCPGC
TTVTSFVFTYPEADSPFSCAAAHRRKSSSNPSDSLSPSTLLAL"

BASE COUNT 512 a 579 c 505 g 511 t
ORIGIN

Alignment Scores:
Pred. No.: 3.2e-13 Length: 2107
Score: 233.00 Matches: 55
Percent Similarity: 65.66% Conservative: 10
Best Local Similarity: 55.56% Mismatches: 16
Query Match: 47.26% Indels: 18
DB: 10 Gaps: 3

US-10-059-720-19 (1-97) x BC029814 (1-2107)

QY 13 SerMetThrGlyGly-----GlnGlnMetGlyArgAspProAspLeuGluGlnGln 30
Db 479 ACCGTGTAGGAGGAGCGGAGCGCATCGGACAG-----AGGGGCAARAGTA 526
QY 31 GluGluLeuAlaArgGluAsnGluGluLysGluAlaGluGluGluGlu 50
Db 527 GAGCAGCTATCTCTGAGAGAGGAGAGAAACGAGAGATCCGAGGACCGATAGATG 586
QY 51 AsnAla-----GluLeuGluAlaGluThr 58
Db 587 GTGTCAGCCCAAGTCCGGAATCGAGAGGGAGCTGACAGATACATACACGCGGAGACA 646

QY 59 AspGlnLeuGluAspGluLeuSerAlaLeuGlnThrGluLeuAlaAsnLeuLeuLysGlu 78
 Db 647 GATCACTTGAAGTGAAGTCTGCTTGCAGACTGAGATTGCCAATCTGCTGAAGAG 706
 QY 79 LysGlnLeuGluPheLeuAlaAlaHisArgProAlaCysLysIleProAsp 97
 Db 707 ARGGAATACTGGAGTTATTTTGGCAGCCCACTGCTGCAAGATCCCAAT 763

RESULT 12
 RNCFOSR RNCFOSR 2116 bp mRNA linear ROD 12-SEP-1993
 LOCUS Rat c-fos mRNA.
 DEFINITION X06769
 ACCESSION X06769.1 GI:59933
 KEYWORDS fos cellular oncogene; fos oncogene; oncogene.
 SOURCE Rattus rattus
 ORGANISM Rattus rattus
 Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.
 1 (bases 1 to 2116)
 Curran, T., Gordon, M.B., Rubino, K.L., and Sambucetti, L.C.
 Isolation and characterization of the c-fos (rat) cDNA and analysis
 of post-translational modification in vitro
 Oncogene 2 (1), 79-84 (1987)
 JOURNAL 88143713
 MEDLINE 3325886
 PUBMED
 COMMENT Data kindly reviewed (05-SEP-1988) by CURRAN T.
 FEATURES
 Location/Qualifiers
 1..2116
 /organism="Rattus rattus"
 /mol_type="mRNA"
 /db_xref="taxon:10117"
 /clone="pcfos(rat)-1"
 /cell_line="PC 12"
 /clone_lib="lambda gt10"
 135..1277
 /note="c-fos protein (AA 1-380)"
 /codon_start=1
 /protein_id="CAA29937.1"
 /db_xref="GI:55934"
 /db_xref="SWISS-PROT:P12841"
 /translation="MFSGFNADYEASSRCSASPAGDSLSYHSPADSPFSGMGPV
 NTQFCADLSSVANSFPTVTAISTPDQLVQPTLVSSVAPSTRAHPYGLPTPS
 TGAYARAGVNTMGGRAQSIGRGVQESPEEEKRIRRRNRKAAKCRNRRE
 LTDLQAEITDLEKALQETIANLKEKLEFLAHRPAKTI PNDLGFPEMSV
 TSLLDTGLPEATPPESEAFPLPLNDPEPKPSLFPVKINSNMLKAPFDPLFPA
 SSRPSGSETARSVPDVLGSGFYAADWEPLHSSSLGMPVTEPLCTPVVCTPSC
 TTYTSSRFVTPPADSPFCNAARHGKSSNSPSSDLSLSPILLAL"

CDS
 275..276
 /note="exon/exon boundary"
 527..528
 /note="exon/exon boundary"
 635..636
 /note="exon/exon boundary"
 1980..2005
 /note="pot. rapid mRNA turnover sequence"
 2092..2097
 /note="polyA signal"
 2116
 /note="polyA site"
 BASE COUNT 524 a 573 c 509 g 510 t
 ORIGIN
 Alignment Scores:
 Pred. NO.: 4e-13 Length: 2116
 Score: 232.00 Matches: 55
 Percent Similarity: 65.66% Conservative: 10
 Best Local Similarity: 55.56% Mismatches: 16
 Query Match: 47.05% Indels: 18
 DB: 10 Gaps: 3

US-10-059-720-19 (1-97) x RNCFOSR (1-2116)
 QY 13 SerMetThrGlyGly-----GlnGlnMetGlyArgAspProAspLeuGluGlnArgAla 30
 Db 474 ACCATGTCAGCGCGCAGAGCGCAGCATCGGAG-----AGGGCAAGTA 521
 QY 31 GluGlnLeuAlaArgGluAsnGluLeuGluLysGluAlaGluGluLeuGluGlnGlu 50
 Db 522 GAGCAGCTATCTCTGAGAGAGAGAGAACGAGAGATCCGAGGGAAGGAATAAGATG 581
 QY 51 AsnAla-----GluLeuGluAlaGluThr 58
 Db 582 CCTGCAGCAAGTCCGGAATCGGAGGAGGAGCTGACAGATACGCTCCAGCGGAGACA 641
 QY 59 AspGlnLeuGluAspGluLysSerAlaLeuGlnThrGluLeuAlaAsnLeuLysGlu 78
 Db 642 GATCACTTGAAGTGAAGTCTGCTTGCAGACTGAGATTGCCAATCTGCTGAAGAG 701
 QY 79 LysGlnLeuGluPheLeuAlaAlaHisArgProAlaCysLysIleProAsp 97
 Db 702 AAGGAATACTGGAGTTATTTTGGCAGCCCACTGCTGCAAGATCCCAAT 758

RESULT 13
 AF033012 1146 bp mRNA linear ROD 09-DEC-1998
 LOCUS Cricetulus griseus c-fos protooncogene protein mRNA, complete cds.
 DEFINITION AF033012
 ACCESSION AF033012
 VERSION AF033012.1 GI:2654078
 KEYWORDS Cricetulus griseus (Chinese hamster)
 SOURCE Cricetulus griseus
 ORGANISM Cricetulus griseus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
 Cricetulus.
 1 (bases 1 to 1146)
 Shyamala, V., Khoja, H., Anderson, M.L., Wang, J.X., Cen, H. and
 Kavanagh, W.M.
 High-throughput screening for ligand-induced c-fos mRNA expression
 by branched DNA assay in Chinese hamster ovary cells
 Anal. Biochem. 266 (1), 140-147 (1999)
 JOURNAL 99105778
 MEDLINE 9887223
 PUBMED
 2 (bases 1 to 1146)
 Khoja, H. and Shyamala, V.
 Direct Submission
 Submitted (04-NOV-1997) Molecular Biology, Chiron Corporation, 4560
 Horton St., Emeryville, CA 94608, USA
 JOURNAL
 FEATURES
 Location/Qualifiers
 1..1146
 /organism="Cricetulus griseus"
 /mol_type="mRNA"
 /db_xref="taxon:10029"
 /cell_lines="CHO"
 /tissue_type="ovary"
 1..1146
 /codon_start=1
 /product="c-fos protooncogene protein"
 /protein_id="AAC83461.1"
 /db_xref="GI:2654079"
 /translation="MFSGFNADYEASSRCSASPAGDSLSYHSPADSPFSGMGPV
 NAQFCADLSSVANSFPTVTAISTPDQLVQPTLVSSVAPSTRAHPYGLPTPS
 TGAYARAGVNTMGGRAQSIGRGVQESPEEEKRIRRRNRKAAKCRNRRE
 LTDLQAEITDLEKALQETIANLKEKLEFLAHRPAKTI PNDLGFPEMSV
 TSLLDTGLPEATPPESEAFPLPLNDPEPKPSLFPVKINSNMLKAPFDPLFPA
 SSRPSGSETARSVPDVLGSGFYAADWEPLHSSSLGMPVTEPLCTPVVCTPSC
 TTYTSSRFVTPPADSPFCNAARHGKSSNSPSSDLSLSPILLAL"

BASE COUNT 255 a 371 c 299 g 221 t
 ORIGIN
 Alignment Scores:
 Pred. NO.: 2.62e-13 Length: 1146
 Score: 231.00 Matches: 55
 Percent Similarity: 64.00% Conservative: 9

```
Best Local Similarity: 55.00% Mismatches: 22
Query Match: 46.86% Indels: 14
DB: 10 Gaps: 2

US-10-059-720-19 (1-97) x AF033012 (1-1146)

QY 12 AlaSerMetThrGlyGlyGlnGlnMetGlyArgAspProAspLeuGluGlnArgAla--- 30
DQ 325 GCGCAATGTAAGACCGTGTGAGCGGCGAGCGGCGAGCATCAGCAGAGCAAA 384
QY 31 ---GluGluLeuAlaArgGluGluGluGluGluGluGluGluGluGluGluGlu 49
DQ 385 GTAGAGCAGTATCTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 444
QY 50 GluAsnAla-----GluLeuGluAlaGlu 57
DQ 445 ATGCTGCAGCCAAATCTCGAATCGGAGGAGGAGTACTGATCTCTCGAGCGGAG 504
QY 58 ThrAspGlnLeuGluAspGluGluGluGluGluGluGluGluGluGluGluGlu 77
DQ 505 ACAGACCAACTTGAAGACGAGAGAGTCTGCACTGAGACTGAGATTCCTCAAA 564
QY 78 GluGluGluGluGluGluGluGluGluGluGluGluGluGluGluGluGluGlu 97
DQ 565 GAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 624

RESULT 14
AF033814 3811 bp RNA linear VRL 03-DEC-1998
LOCUS Murine osteosarcoma virus, complete genome.
DEFINITION AF033814
ACCESSION AF033814.1 GI:2801477
VERSION Murine osteosarcoma virus
KEYWORDS Murine osteosarcoma virus
SOURCE Murine osteosarcoma virus
ORGANISM Viruses; Retroviral taxonomy, protein structure, sequences, and
retroviruses; 1-Mammalian type C virus group.
REFERENCE 1 (bases 1 to 3811)
AUTHORS Petropoulos, C.J.
TITLE Appendix 2: Retroviral taxonomy, protein structure, sequences, and
genetic maps
JOURNAL (in) Coffin, J.M. (Ed.):
RETROVIRUSES: 757;
Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York,
NY, USA (1997)
REFERENCE 2 (bases 1 to 3811)
AUTHORS Chappay, C.
TITLE Direct Submission
JOURNAL Submitted (12-NOV-1997) NIH, NLM, Rockville Pike, Bethesda, MD
20894, USA

FEATURES
source Location/Qualifiers
1..3811
/organism="Murine osteosarcoma virus"
/mol_type="genomic RNA"
/db_xref="taxon:11830"
69..3217
/gene="gag-FOS"
69..3217
/gene="gag-FOS"
69..145
/product="p75"
69..145
/gene="gag-FOS"
637..2301
/gene="gag-FOS"
637..2301
/codon_start=1
/product="p75"
/db_xref="GI:2801478"
/db_xref="AF033814"
/translation="MGGTATLPLSLTLEHWDQVQASQSDVVKRRVWVTCSEAWP
TDDGFMQDGLNLDLILQKSKVSPGPHGPDQVYIVTWETATYEPVPPVPPVPS
PKLSPGPTAPLPSGPTQPPRSALYPAITSIKPRSPQVLDGGLDILDLTET
PPYGGQSPSDGSDGDEATSTSEIPAPSPVSRGRKRPDPAADTSFAFLRL
CGNGQKNNPSPSFDGPKLTALIESVLTTHQPTWDDCQQLLTLTGEBKQVLEAR"

mat_peptide /gene="gag-FOS"
mat_peptide /product="MA"
mat_peptide /gene="gag-FOS"
mat_peptide /product="CA"
mat_peptide /gene="gag-FOS"
mat_peptide /product="FOS"
3'UTR /gene="gag-FOS"
BASE COUNT 888 a 1094 c 968 g 861 t
ORIGIN

Alignment Scores:
Pred. No.: 9,23e-13 Length: 3811
Score: 231.00 Matches: 56
Percent Similarity: 63.47% Conservative: 10
Best Local Similarity: 58.95% Mismatches: 19
Query Match: 46.86% Indels: 10
DB: 14 Gaps: 4

US-10-059-720-19 (1-97) x AF033814 (1-3811)
QY 13 SerMetThrGlyGly-----GlnGlnMetGlyArgAspProAspLeuGluGlnArgAla 30
DQ 1834 ACCGTCTGAGGAGCGAGCGGAGCATCGGAGAGGCGGAGAGAGAGAGAGAGATCT 1893
QY 31 -----GluGluLeuAlaArg-----GluAsnGluGluGluGluGluGluGluGlu 46
DQ 1894 CTGAGAGAGAGAGTGAACGAGAGATCCGAGAGAGAGAGAGAGATGCTCGAGCCAG 1953
QY 47 LeuGluGlnGluAsnAlaGly-----LeuGluAlaGluThrAspGlnLeuGlu 62
DQ 1954 TCCCGGATCGGAGGAGGAGGCTGACAGATACCTCCAGCGGAGAGAGAGATCACTTGA 2013
QY 63 AspGluGluSerAlaLeuGlnThrGluLeuAlaAsnLeuLeuGluGluGluGluGlu 82
DQ 2014 GATGAGAGATCTGCGTTCGAGACTGAGATTCGCAATCTGCTGAGAGAGAGAGAGATG 2073
QY 83 GluPheLeuLeuAlaAlaHisArgProAlaCysLysLysLeuProAsp 97
DQ 2074 GAGTTATTTTGGCAGCCCGACCTGCTGCAAGATCCCGAT 2118

RESULT 15
MSVUSV
LOCUS 3811 bp DNA linear VRL 14-FEB-2003
DEFINITION FBR murine osteosarcoma, complete proviral sequence integrated in
Rattus norvegicus genome.
ACCESSION K02712 X03347
VERSION K02712.2 GI:28374460
KEYWORDS FBR murine osteosarcoma virus
ORGANISM FBR murine osteosarcoma virus
VIRUSES; Retroviral taxonomy, protein structure, sequences, and
retroviruses; 1-Mammalian type C virus group.
REFERENCE 1 (bases 1 to 3811)
AUTHORS Van Beveren, C., Enami, S., Curran, T. and Verma, I.M.
TITLE FBR murine osteosarcoma virus. II. Nucleotide sequence of the
provirus reveals that the genome contains sequences acquired from
two cellular genes
JOURNAL Virology 135 (1), 229-243 (1984)
MEDLINE 84225828
PUBMED 6203215
REFERENCE 2 (bases 1 to 3811)
AUTHORS Van Beveren, C., Enami, S., Curran, T. and Verma, I.M.
TITLE Direct Submission
JOURNAL Submitted (01-JUL-1986) The Salk Institute, La Jolla, CA, USA
```

